

FIGURE 1

ID-65

Clone 3-60

5
GTGTTTATGATGAAAAAAGGACAAGTAAATGATACTAAGCAA
TCTTACTCTCTACGTAAATATAAATTTGGTTTAGCATCAGTAA
TTTTAGGGTCATTCATAATGGTCACAAGTCCTGTTTTTGCGGA
TCAAACCTACATCGGTTCAAGTTAATAATCAGACAGGCACTAG
10 TGTGGATGCTAATAATTCTTCCAATGAGACAAGTGCGTCAAGT
GTGATTACTTCCAATAATGATAGTGTTCAAGCGTCTGATAAAG
TTGTAAATAGTCAAAATACGGCAACAAAGGACATTACTACTC
CTTTAGTAGAGACAAAGCCAATGGTGGAAAAAACATTACCTG
AACAAGGGAATTATGTTTATAGCAAAGAAACCGAGGTGAAAA
15 ATACACCTTCAAAATCAGCCCCAGTAGCTTTCTATGCAAAGAA
AGGTGATAAAGTTTTCTATGACCAAGTATTTAATAAAGATAAT
GTGAAATGGATTTTCATATAAGTCTTTTGGTGGCGTACGTCGAT
ACGCAGCTATTGAGTCACTAGATCCATCAGGAGGTTTCAGAGA
CTAAAGCACCTACTCCTGTAACAAATTCAGGAAGCAATAATC
20 AAGAGAAAATAGCAACGCAAGGAAATTATACATTTTTCACATA
AAGTAGAAGTAAAAAATGAAGCTAAGGTAGCGAGTCCAATC
AATTTACATTGGACAAAGGAGACAGAATTTTTTACGACCAA
TACTAACTATTGAAGGAAATCAGTGGTTATCTTATAAATCATT
CAATGGTGTTTCGTCGTTTTGTTTTGCTAGGTAAAGCATCTTCA
25 GTAGAAAAAACTGAAGATAAAGAAAAAGTGTCTCCTCAACCA
CAAGCCCGTATTACTAAAACCTGGTAGACTGACTATTTCTAACG
AAACAACCTACAGGTTTTGATATTTTAATTACGAATATTAAAGA
TGATAACGGTATCGCTGCTGTTAAGGTACCGGTTTGGACTGAA
CAAGGAGGGCAAGATGATATTAATGGTATACAGCTGTAACCT
30 ACTGGGGATGGCAACTACAAAGTAGCTGTATCATTGCTGAC
CATAAGAATGAGAAGGGTCTTTATAATATTCATTATATACTACC
AAGAAGCTAGTGGGACACTTGTAGGTGTAACAGGAACTAAAG
TGACAGTAGCTGGAACCTAATTCTTCTCAAGAACCTATTGAAAA
TGGTTTACCAAAGACTGGTGGTTTATAATATTATCGGAAGTACT
35 GAAGTAAAAAATGAAGCTAAAATATCAAGTCAGACCCAATTT
ACTTTAGAAAAAGGTGACAAAATAAATTATGATCAAGTATTG
ACAGCAGATGGTTACCAGTGGATTTCTTACAAATCTTATAGTG
GTGTTTCGTCGCTATATTCCTGTGAAAAAGCTAACTACAAGTAG
TGAAAAAGCGAAAGATGAGGCGACTAAACCGACTAGTTATCC
40 CAACTTACCTAAAACAGGTACCTATACATTTACTAAAACCTGTA
GATGTGAAAAGTCAACCTAAAGTATCAAGTCCAGTGGAATTT
AATTTTCAAAGGGTGAAAAAATACATTATGATCAAGTGTTA
GTAGTAGATGGTCATCAGTGGATTTTCATACAAGAGTTATTCCG
GTATTCGTCGCTATATTGAAATTTAA

15

20

- ID-66

Clone 3-5

ATGATATTGAGACGTGCAACTATTGTTTTATGGCAACTGGGTATCGCCATT
TCTCTCATTCTTAGTATTCTAGCCTTAAATCTTTATTTCCATAGTACTCCCTT
30 GCAAACCAATGCAGCTTTACGGAACCTTGCTCCTTCATTAAACCATCTTTTT
GGGACAGATGGTTTAGGTAGGGATATGTTTGTGACAACGATTAAAGGACT
TTATTTCTCTCTACAAGTCGGCTTATTAGGTGCCCTTATGGGGGTCATTCTG
GCGACAGTTTTTTGGAGTGCTTGCAGGTTTAGGAAATAGCATTATTGATAAA
ATAATAGCATGGTTAGTTGATTTGTTTATTGGTATGCCTCATTGATTTTTA
35 TGATTCTCATTTCTTTTGTTGTTGGGAAAGGTGCTCAAGGGGTCATCATTGC
AACGGCTGTTACACATTGGCCTTCTTTAGCAAGGCTTATCCGCAATGAAGT
CTATCATCTAAAGAATAAAGAATTTGTCCAACCTTCTAAAAGTATGGGAAA
AACGCCTTATTATATTGTGAGGCATCATATCCTGCCTTTGATTGCTTCTCAA
ATTTTCATTGGTTTTATCCTCTTATTTCCACATGTCATCCTACATGAAGCAT
40 CAATGACTTTCTTAGGATTTGGGCTCTCTGCCGAACAACCTTCGGTTGGTA
TCATTCTGTCAGAGGCAGCTAAGCATATCTCTCTTGGAATTTGGTGGTTGG
TTATCTTTCCAGGACTTTATCTTATTTTGGTTGTCAATGCATTTGATACTAT
CGGAGAATCTTTAAAGAACTCTTTTACCCTCAAACCTGATCATTTTTTAG

MILRRRTIVLWQLGIAISLILSILALNLYFHSTPLQTNAALRNLAPSLNHLFGTD
GLGRDMFVRTIKGLYFSLQVGLLGALMGVILATVFGVLAGLGNSIIDKIIAWL
VDLFIGMPHLIFMILISFVVGKGAQGVIIATAVTHWPSLARLIRNEVYHLKNKE
5 FVQLSKSMGKTPYYIVRRHHILPLIASQIFIGFILLFPHVILHEASMTFLGFGLSAE
QPSVGIILSEAAKHISLGNWWLVIFPGLYLILVVNAFDTIGESLKKLFPQTDHF
*

10 Sequence description

A) Length: 822 bp - 274 aa (full length gene)

B) Sequence Characteristics:

Potential leader peptide sequence

15 Orf is preceded by a potential Shine-
Dalgarno sequence.

ID-78

20

Clone 3-5b

ATGACAGAAACATTATTAAGCATTAAAGACCTCTCCATCACCTTCACTCAA
TACGGAAGATTTTTTAAAACCATTTCAATCAACACCGATACAAGCGCTGA
25 ATTTAGAAATTA AAAAAGGTGAGTTATTAGCTATTATAGGTGCTAGTGGTT
CGGGGAAGAGTTTATTAGCACATGCTATTATGGATATTCTTCCTAAAAATG
CATCTGTAACAGGAGATATGATTTATCGTGGTCAATCACTAAATTCTAAAC
GCATTAAACAGTTGCGAGGAAAAGATATTACGTTGATTCCACAATCAGTTA
ATTATTTAGATCCATCTATGAAAGTCAAACATCAGGTGCGCTTAGGTATCT
30 CAGAAAATTCAAAGGCTACTCAAGAAGGATTGTTTCAACAGTTTGGTTTAA
AAGAAAGTGATGGTGACTTGGATCCTTTCCAACCTTTCTGGCGGAATGCTCC
GACGTGTTTTGTTTACAACGTGTATTAGTGATAAGGTTTCTTTGATTATTGC
GGATGAGCCCACCCCTGGATTACATCCAGATGCTCTGCAAATGGTTTTAGA
CCAACCTACGCTCCTTTGCAGATAAAGGAATAAGCGTTATATTTATCACTCA
35 TGATATTGTAGCAGCTAGTCAAATTGCTGATCGTATTACTATTTTTAAAGA
GGGAAAAGCTATTGAAACAGCTCCAGCTAGTTTCTTTAGCGGAAATGGAG
AGCAGTTACAAACAGAATTTGCTAGAAGTTTATGGCGCTCTCTCCACAGC
AAGAATTTTTGAAAGGAGTTACTCATGACCTTAGAGGCTAA

40

MTETLLSIKDLSITFTQYGRFLKPFQSTPIQALNLEIKKGELLAHIGASGSGKSL
AHAIMDILPKNASVTGDMIYRGQSLNSKRIKQLRGKDITLIPQSVNYLDPSMK
VKHQVRLGISENSKATQEGLFQQFGLKESDGDLDPFQLSGGMLRRVLFTTCIS
DKVSLIIADEPTPGLHPDALQMVLDQLRSFADKGISVIFITHDIVAASQIADRITI
FKEGKAIETAPASFFSGNGEQLQTEFARSLWRSLPQQEFLKGVTHDLRG*

5

10

15

20

25

35

40

This gene was not isolated using the LEEP system. However in determining a full length gene sequence for ID-76, this gene was identified upstream.

ID-80

5

Clone 2-17

TTGCGGACAATTACGTTCAAACACAATGAAACGCGATCGTCAAAAAGCGA
AGGTAGGGCGGTAATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGA
10 CAAAGATTAGTCGTCGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTAT
TATTGTCAGGATGTGTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAAC
TCCGAATGTTTCAGCGCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTAA
TTCTTTTCGTAGTTTAGATCAGCTAACTAGCTTTAAAGAGATTTTTTGGGTT
ATTGGTCAAAATGTAGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGT
15 TACTATCCCTAAAGCCAAGTTTACGGAAATATAAAAGCGTTATATTACTTG
CTTTCTTGATGTCTCTTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTT
AATAGATGCTAATCGGGTTTTTGAATCGACGATCTATGGACAAATACCTT
AGGCGGTCCTTTCCGCCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCT
AACTATTAGAAAATGA

20

MRTITFKHNETRSSKSEGRAVMLKRLFTEDGELTKISRRFVWMLVVIYCLIIVR
MCFGPQIMIEGVSTPNVQRFGRIVALLVPFNSFRSLDQLTSFKEIFWVIGQNVV
NILLFPLIIGLLSLKPSLRKYKSVILLAFLMSLFIECTQVVLIDILIDANRVFEIDD
LWTNTLGGPFALWSYRNIKGWLLTIRK*

25

Sequence description

- 30 A) Length: 579 bp - 193 aa (full length gene)
B) Sequence Characteristics:
Possesses a potential leader peptide sequence
No obvious Shine-Dalgarno, but the 'TTG' codon
may not be the actual translation start point.
A methionine (ATG) that occurs ~22 codons
35 downstream of the 'TTG' is preceded by a
potential Shine-Dalgarno sequence and may
represent the actual start codon.

ID 81

40

Clone 3-1

TTGAAAAATTTAAATCGTTATGTAGTTGCGGTTTCTGGAGTCGTTTTACATT
 TAATGCTAGGATCAACTTATGCTTGGAGTGTGTTTCGTAACCCAATTATCT
 CAGAGACTGGTTGGGATATTTTCATCAGTTTCATTTCGCTTTTAGTTTGGCTAT
 TTTTGTCTAGGAATGTCTGCAGCTTTTATGGGACACTTAGTAGAGCGTTTT
 5 GGTCCTAGGATAATGGGAATGATTTCTGCTATTTTATATGGAGCAGGGAAT
 GTGTTAACAGGCTTAGCCATTGAACTCAGCAGTTATGGTTACTGTATGTT
 GCATACGGTATTTTAGGAGGAATCGGACTTGGTTCAGGTTATATTACTCCA
 GTATCGACTATTATTAATGGTTTCCTGATAGGAGGGGACTAGCAACAGG
 ATTCGCTATTATGGGATTTGGCTTTGCTTCTTTAGTAACAAGTCCGCTTGCA
 10 CAATCCTTACTGATTAGGATTGGTGTGGGTAAAACGTTTTATATTTTGGGA
 TTAGTATATTTTTTTGTCATGATGATTGCCTCACAATTTATTAACAACCAC
 CTCAGGAAAAAATAACTATTTTGACTCACGATGGTAAAAAGAATGCTATG
 AATTCACAAATTATCACTGGATTAAAAGCAAACGTCGCTATAAAATCAAA
 AACCTTTTACATCATTGTTGACCTTGTTTATTAATATTTTCGTGTGGCTTA
 15 GGTTTAATATCAGCAGCTTCACCAATGGCACAAGATTTAGCAGGCTATTCC
 GCAGAATCTGCAGCCTTATTAGTAGGGGTACTAGGGATATTTAACGGTTTT
 GGACGTCTGTTATGGGCAAGTCTCTCTGACTACATTGGACGCCCGTTGACC
 TTTATAATATTATTTATTGTGAACCTTATTATGACTTCTAGTTTATTTTTGTC
 ATTC AATGCTATTGTATTTGCAATAGCGATGTCTATTTTAATGACTTGTTAT
 20 GGTGCAGGTTTTTCCTTATTACCTGCTTATCTAAGTGATATTTTGGAAACAA
 AGGAATTAGCTACTTTACATGGTTATAGTTTAACAGCATGGGCAATAGCAG
 GTCTGTTTGGGCCCTATTGTTATCAAAGACATATTCATGGGGAAATTCCT
 ATCAATTGACATTAATGGTTTTTGGTTTTTTATTCTTATTCGGATTATTGTTA
 TCTCTATATTTAAGAAAATTAACAACATAAGTTGTGTAG
 25 LKNLNRYVVA VSGVVLHMLGSTYAWSVFRNPIISETGWDISSVSFAFSLAIFC
 LGMSAAFMGHLVERFGPRIMGMISAILYGAGNVLTGLAIETQQLWLLYVAYG
 ILGGIGLGSYITPVSTIIKWFPDRRLATGFAIMGFGFASLVTSPLAQSLIRIG
 30 VGKTFYILGLVYFFVMMIASQFIKQPPQEKITILTHDGKKNAMNSQIITGLKAN
 VAIKSKTFYIIWLTFLFINISGLGLISAASPM AQDLAGYSAESAALLVGVLGIFN
 GFGRLLWASLSDYIGRPLTFIILFIVNFIMTSSLFSLFNAIVFAIAMSILMTCYGA
 GFSLLPAYLSDIFGTKE LATLHGYSLTAWAIAGLFGPLLLSKTYSWGN SYQLTL
 MVFGFLFLFGLLLSLYLRKLTTKVV*

35

Sequence description:

40

- A] Length 1221 bp - 407 a.a (full length gene).
- B] TTG start codon with Shine-Dalgarno sequence upstream. Obvious signal peptide, with hydropathy plot exhibiting many possible membrane spanning regions, indicating protein to be transmembrane.

5 ID-82

Clone 48

10 ATGGCAGATAAAAAACAGAACATTTAAACTTGTAGGTGCAGGATCTTCTAG
CACACAAGAAAAAATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTG
GCGTCGCTTGAAAAAAAACAAATTAGCAGTAGTTTCACTCTATTTATTAGC
TCTTTTACTTACTTTTTTCGTTAGCCTCAAATTTATTTGTAACCTCAGAAGGAT
15 GCTAATGGGTTTGATTTCGAAAAAAGTAACGACATATCGCAACTTACCACCT
AAATTGAGTTCAAACCTTCCTTTTTTGAATGGTAGCATTAAATCCATCA

MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLA
LLLTFSLASNLFVTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSINPS

20

Sequence description:

A] Current length is 303 bp - 101 aa
B] No obvious signal peptide but Shine
25 Dalgarno sequence upstream of the ATG start
codon. Not identified directly using the LEEP system but was found
directly downstream of ID-34 described in WO 00/06736.

30 ID-83

Clone 98

35

ATGAAAATAGTAGTACCAGTAATGCCTCGCAGTCTTGAAGAGGCTCAAGA
AATAGATTTATCAAAATTTGATAGTGTTGATATTATTGAATGGCGAGCTGA
TGCCTTACCAAAGGATGACATTATTAATGTAGCTCCAGCTATTTTTGAGAA
ATTCGCAGGTCATGAAATTATTTTTACTTTTCGTACAACGCGTGAAGGTGG
40 TAATATTGTCTTATCTGATGCTGAGTATGTTGAGTTAATCCAGAAAATTAA
TTCTATCTACAATCCAGATTATATTGATTTTGAGTATTTTTCACATAAAGAA
GTTTTTCAAGAAATGCTAGAATTTCCAAATTTAGTCCTGTCTTATCACAATT
TTCAAGAGACACCGGAGAATATTATGGAGATATTTTCAGAAATTAACAGCC
CTAGCACCACGAGTTGTGAAAATCGCAGTAATGCCAAAGAATGAACAAGA

5 TGTCTTAGACGTTATGAATTACACTCGCGGTTTCAAGACTATTAATCCTGA
TCAAGTTTATGCGACGGTATCTATGAGTAAAATTGGACGTATTTCTCGTTTT
GCTGGTGATGTAACCTGGATCTAGTTGGACATTTCATATTTAGATTCATCT
ATCGCACCCGGACAAATTACTATTTTCAGAGATGAAGCGTGTCAAAGCATT
GCTTGACGCTGACTGA

10 MKIVVPVMPRSLEEAQEIDLSKFDSVDIIEWRADALPKDDIINVAPAIFEFK
HEIIFTFRTTREGGNIVLSDAEYVELIQKINSIYNPDYIDFEYFSHKEVFQEMLEF
PNLVLSYHNFQETPENIMEIFSELTALAPRVVKIAVMPKNEQDVLDMNYTRG
FKTINPDQVYATVSMISKIGRISRFAGDVTGSSWTFAYLDSSIAPGQITISEMKRV
KALLDAD*

15 Sequence description:

- A] Length 678 bp, 225 aa (full length gene)
- B] No obvious signal peptide, but there is a
Shine Dalgarno immediately upstream of ORF.

20 ID-84

25 Clone RS-52

30 ATGAAAGACTTATTTGCAACAACAGAAGCATCATCAAGGAAACAGGAACA
AGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAACAAATGGTA
ATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAATGAGATGACA
GGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCGATTAGTTTT
TCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATATCAACCGAA
TGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCAAATGTTAATGA
TATTGAAGTGATTTATATGAAGAAAGAATAG

35 MKDLFATTEASSRKQEQRIVNYIKQHVELTNGNQIKKIEFIDFQKNEMTGTW
GISTKINEQFSISFSEDRIGGKLRLALGYQPNEIGFSKDINSNNQNVNDIEVIYMK
KE*

40 Sequence description:

- A] length: 333 bp - 111 aa (partial sequence)
- B] No obvious Shine Dalgarno sequence upstream
of the ATG start codon, and no obvious signal
peptide within the protein.

ID-85

5

Clone RS-53

10 ATGAAAAAACGTATATGGTATTTGATAATAATAATCACAGTAATTTTAGGA
GGACTAGCCATGAAAAACTTATTTGCAACAACAGAAGCATCATCAAGGAA
ACAGGAACAAGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAA
CAAATGGTAATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAT
GAGATGACAGGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCG
15 ATTAGTTTTTCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATAT
CAACCGAATGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCA

20 MKKRIWYLIHITVILGGLAMKNLFATTEASSRKQEQDRIVNYIKQHVELTNGN
QIKKIEFIDFQKNEMTGTWGISTKINEQFSISFSEDRI GGKLRALGYQPNEIGFSK
DINSNNQ

Sequence description:

25 A] Length: 351 bp - 117 aa (Partial sequence)
B] Obvious signal peptide and Shine Dalgarno
sequence upstream of the ATG start codon.

30

ID-86

Clone ID-74

35 ATGTCAAATCAATATGATTATATCGTTATTGGTGGAGGTAGT
GCAGGCAGTGGTACCGCTAATAGGGCAGCCATGTATGGAGC
AAAAGTCCTGTTAATTGAAGGTGGACAAGTAGGTGGAAGTTG
TGTAACTTAGGTTGTGTACCTAAGAAAATCATGTGGTATGG
TGCACAAGTTTCTGAGACACTCCATAAGTATAGTTCAGGTTA
40 TGGTTTTGAAGCCAATAATCTTAGTTTTGATTTTACTACTCTA
AAAGCTAATCGCGATGCTTACGTGCAGCGGTCTAGACAGTCG
TATGCCGCTAATTTTGAGCGTAATGGGGTCGAAAAGATTGAT
GGATTGCTCGTTTTATTGATAACCATACTATTGAAGTGAATG
GTCAGCAATATAAAGCTCCTCACATTACTATTGCAACAGGTG

GACACCCTCTTTACCCTGATATTATTGGAAGTGAACCTTGGTG
AGACTTCTGATGATTTTTTTGGATGGGAGACCTTACCAAATTC
TATATTGATTGTTGGGGCGGGCTATATCGCGGCAGAACTTGC
TGGAGTGGTTAATGAATTAGGCGTTGAAACCCATCTTGCATT
5 TAGAAAAGACCATATTCTACGCGGATTTGATGACATGGTAAC
AAGTGAGGTTATGGCTGAAATGGAGAAATCAGGTATCTCTTT
ACATGCTAACCATGTACCTAAATCTCTTAAACGCGATGAAGG
TGGCAAGTTGATTTTTGAAGCTGAAAATGGGAAAACGCTTGT
CGTTGATCGTGTAATATGGGCTATCGGCCGTGGACCAAATGT
10 AGACATGGGACTTGAAAATACCGATATTGTTTTAAATGATAA
AGATTATATCAAAACAGATGAATTTGAGAATACTTCTGTAGA
TGGCGTGTATGCTATTGGAGATGTTAATGGGAAAATTGCCTT
GACACCGGTAGCAATTGCAGCAGGTCGTCGCTTATCAGAAAG
ACTTTTTAATCATAAAGATAACGAAAAATTAGATTACCATAA
15 TGTACCTTCAGTTATTTTTACTCACCTGTAATTGGGACGGTA
GGACTTTCAGAAGCAGCAGCTATCGAGCAATTTGGAAAAGAT
AATATCAAAGTCTATACATCAACTTTTACCTCTATGTATACGG
CTGTTACCAGTAATCGCCAAGCAGTTAAGATGAAGCTCATAA
CCCTAGGAAAAGAGGAAAAAGTTATTGGGCTTCATGGTGTTG
20 GTTATGGTATTGATGAAATGATTCAAGGTTTTTCAGTTGCTAT
CAAAATGGGGGCTACTAAAGCAGACTTTGATGATACTGTTGC
TATTCACCCAACCTGGATCTGAGGAATTTGTTACAATGCGCTA
A
25 MSNQYDYIVIGGGSAGSGTANRAAMYGAKVLLIEGGQVGGTC
VNLGCVPPKIMWYGAQVSETLHKYSSGYGFEANNLSFDFTTLK
ANRDAYVQSRQSYAANFERNGVEKIDGFARFIDNHTIEVNGQ
QYKAPHITATGGHPLYPDIIGSELGETSDDFFGWETLPNSILIVG
AGYIAAELAGVVNELGVETHLAFRKDHILRGFDDMTSEVMAE
30 MEKSGISLHANHVPKSLKRDEGGKLIFEANGKTLVVDRVIWAI
GRGPNVDMGLENTDIVLNDKDYIKTDEFENTSVDGVYAIGDVN
GKIALTPVAIAAGRRLSERLFNHKDNEKLDYHNVPSVIFTHPVIG
TVGLSEAAAIEQFGKDNIVYTSTFTSMYTAVTSNRQAVKMKLI
TLGKEEKVIGLHGVGYGIDEMIQGFSAIKMGATKADFDDTVAI
35 HPTGSEEFVTMR*

40

ID-87

Clone RS-55

5 ATGACAAAAAACATCTTAAAACGCTTGCCTTGGCACTTACTACAGTATCA
GTAGTGACATACAGCCAGGAGGTATATGGATTAGAAAGAGAGGAATCGGT
CAAACAAGAACAAACCCAGTCAGCTTCAGAAGATGATTGGTTCGAAGAAG
ATAATGAGAGGAAAAACAAATGTTTCTAAAGAGAATTCTACTGTTGATGAA
ACAGTTAGTGATTTATTTTCTGATGGAAATAGTAATAACTCTAGTTCTAAA
ACCGAGTCAGTGGTAAGTGACCCTAAACAAGTCCCCAAAGCAAAACCAGA
10 GGTTACACAAGAAGCAAGCAATTCTAGTAATGATGCTAGCAAAGTAGAAG
TACCAAAACAGGATACAGCTTCAAAAAAGGAAACTCTAGAAACATCAACT
TGGGAGGCAAAAGATTTTCGTAAGTACTAGAGGGGATACTTTAGTAGGTTTTTCA
AAATCTGGAATTAATAAGTTATCTCAAACATCACACTTGGTTTTTACCAAGT
CATGCAGCAGATGGAAGTCAATTGACACAAGTAGCTAGCTTTGCTTTTACT
15 CCAGATAAAAAGACGGCCATTGCAGAATATACAAGTAGGCTAGGAGAAA
ATGGGAAACCGAGTCGTTTAGATATTGATCAGAAGGAAATTATTGATGAG
GGAGAAATATTTAATGCTTACCAGTTGACTAAGCTTACTATTCCAAATGGT
TATAAGTCTATTGGTCAAGATGCTTTTGTGGACAATAAGAATATTGCTGAG
GTTAACCTTCCTGAGAGTCTCGAGACTATTTTCAGACTATGCTTTTGCTCACA
20 TGTCTTTAAACAAGTAAAGTTACCAGATAACCTAAAGGTCATTGGAGAA
TTAGCTTTTTTTGATAATCAGATTGGTGGTAAGCTTTACTTGCCACGTCCT
TGATAAAATTAGCAGAACGCGCTTTCAAATCTAATCGTATTCAAACAGTTG
AATTTTTGGGAAGTAAGCTTAAGGTTATAGGAGAAGCAAGTTTTCAAGAT
AATAATCTGAGGAATGTTATGCTTCCGGATGGACTTGAAAAAATAGAATC
25 AGAAGCTTTTACAGGAAATCCAGGAGATGAACATTACAACAATCAGGTTG
TATTGCGCACAAGGACAGGCCAAAATCCACATCAACTTGCGACTGAGAAT
ACTTACGTCAATCCGGACAAATCATTGTGGCGTGCAACACCTGATATGGAT
TATACCAAATGGTTAGAGGAAGATTTTACCTATCAAAAAAATAGTGTTACA
GGTTTTTCAAATAAAGGCTTACAAAAGGTAAGACGTAATAAAAACTTAGA
30 AATTCCAAAACAACACAATGGTATTACTATTACTGAAATTGGTGATAACGC
TTTTCGCAATGTTGATTTTCAAAGTAAACTTTACGTAAATATGATTTGGA
AGAAATAAAGCTCCCCTCAACTATTCGGAAAATAGGTGCTTTTTGCTTTTCA
ATCTAATAACTTGAAATCCTTTGAAGCAAGTGAAGATTTAGAAGAGATTA
AAGAGGGAGCCTTTATGAATAATCGTATTGGAAGTCTAGACTTGAAAGAC
35 AAAGTTATCAAAATAGGTGATGCTGCTTTCCATATTAATCATATTTATGCC
ATTGTTCTTCCAGAATCTGTACAAGAAATAGGACGTTTCAGCTTTTCGACAA
AATGGTGCGCTTCACCTTATGTTTATCGGAAATAAGGTTAAACAATTGGT
GAAATGGCTTTTTTATCCAATAAACTGGAAAGTGTAATCTCTCTGAGCAA
AAACAATTAAAGACAATTGAGGTCCAAGCTTTTTTCGGATAATGCCCTTAGT
40 GAAGTAGTCTTACCGCCAAATTTACAGACTATTCGTGAAGAGGCTTTCAA
AGGAATCATTGAAAGAAGTGAAGGGTTCATCTACATTATCTCAGATTACT
TTTAATGCTTTTGATCAAAATGATGGGGACAAACGCTTTGGTAAGAAAGTG
GTTGTTAGGACACATAATAATTCTCATATGTTAGCAGATGGTGAGCGTTTT
ATCATTGATCCAGATAAGCTATCTTCTACAATGGTAGACCTTGAAAAGTT

15

40

A] Length 3168 bp - 1056 aa (Partial sequence)
B] Obvious signal peptide with Shine Dalgarno sequence upstream of the ATG start codon.

Clone RS-56

AGYIMHKHEAIVSCWGQPRKTCRHKLKISLQSVHNKETGKSAFNDKERLAI

A] Length:153 bp - 51 aa (partial sequence)
B] No signal peptide visible, insufficient
sequence data to determine the presence of a
Shine Dalgarno sequence.

Clone RS-58

40 GTGTCAATTATGCAAAGAAAATCCTATTTAAAATCCATGAGTGTTCCTTACT
TTAACAGCTTGTCTTATATCAGGATATGTGGTTAAAGATATTGCTATGTTA
CATGCAGTATCTGCCAGTGAGAAGAAAGCAAATAATGTCAGTCCGAGAGA
AAATCTCTACAGGGCTGTCAATGATAATTGGCTAGCCAATACAAAACCTCA
AACAAAGGGCAGACTAGTGTTAATAGTTTTTCAGAAATTGAGGATAAATTA
AAGCAACTGTTAGTGTCTGATATGGCTAAAATGGCCTCAGGAAAGATTGA

5 AACAAACCAATGATGAACAGAAAAAATGGTTGCATACTATAAAACAAGGTA
TGGACTTTAAAACAAGAGATAAAAAATGGTCTCAAACCTCTAAAACCAGTT
TTACAAAAAAGTTGAAGCAGTCTCTTCAATGAAAGACTTTCAAAGTTTGGCC
CATGATTTTGTGATGAGTGGTTTTGTTTTACCATTGTTTGGTTTGAAGTGA
10 CCAATGCTCGAGATAATAGCCAAAAGCAATTGGTGCTTCGTCAAGCACCC
GCATTACTTGAATCACCTGACCAATATAAGAAGGGCAATAAAGAAGGTGA
GGCTAAATTATCAGCTTACCGTACTTCAGCAATGGCTTTGCTTAAACAAGC
TGGAAAAAGTAACATTGAAGATAGAAAACTAGTTAAACAAGCTATAGCAT
TTGATAGACTCTTATCAGAAAAAACGCAAGTTGATCAAAGTAAAATCACA
15 GCTGAAAGTGAGACAGCTGCGGGGCGATATAACCCTGAAAGTATGGAAAC
GGTTCACAATTACGCCAAGGAATTTGACTTTAAAGAATTGATTGAAAAACT
AGTTGGGCCAACGAATAAGGCAGTCAATGTAGAAGATAAAAAGTTATTTTA
AACAGGTTAATGATGTTATAAATAGTAAACAATTAGCCAATATGAAAGCA
TGGATGATGATTTCTATGCTAGTTGATCAATCAGATTTTCTAGGAGAACAA
20 AATCGTCAAGCAGCGAGTGCTTTTAAGAATGTTGCGTCTGGTTTGAAGTCA
ATTGAATCGAAAGAAAAAATGCTTACACCCAATTAG

25 MSFMQRKSYLKSM SVLTLTACLISGYVVKDIAM LHAVSASEKKANNVSPREN
LYRAVNDNWLANTKLKQGQTSVNSFSEIEDKLKQLLVSDMAKMASGKIETTN
DEQKKMVAYYKQGMDFKTRDKNGLKPLKPVLQKLEAVSSMKDFQSLAHDF
VMSGFVLPFGLTVETNARDNSQKQLVLRQAPALLES PDQYKKGNKEGEAKLS
30 AYRTSAMALLKQAGKSNIEDRKL VKQAI AFDRLLSEKTQVDQSKITAESETAA
GRYNPESMETVHNYAKEFD FKELIEKL VGPTNKAVNVEDKTYFKQVNDVINS
KQLANMKAWMMISMLVDQSDFLGEQNRQAASAFKNVASGLTQIESKEKMLT
PN*

35 Sequence description:

35 A] Length: 1095 bp - 365 aa (full length gene)
B] an GTG (possible ATG start codon located 7 bp
further downstream) start codon with an obvious
signal peptide. Shine Dalgarno sequence present
upstream of the ORF.

40 ID-90

Clone RS-59

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTTAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTTAAATATTTTTTATAG

5

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLFK
YFL*

10 Sequence description:

A] Length: 174 bp - 58 aa(full length gene)
B] No obvious signal peptide, but Shine
Dalgarno sequence is present upstream of ATG
start codon.

15

20 ID-91

Clone RS-62 (partial sequence)

25

ATGCAGGTATTTTTAAATATTGTCAATAAATTCTTTGATCCAGTTATTCATA
TGGGTTCGGGAGTTGTGATGCTAATTGTCATGACAGGTTTAGCCATGATAT
TTGGAGTGAAGTTTTCTAAAGCACTTGAAGGTGGTAT

30

MQVFLNIVNKFFDPVIHMGSGVVMLIVMTGLAMIFGVKFSKALEGG

Sequence description:

35

A] Length:141 bp - 41 aa (partial sequence
B] Shine Dalgarno sequence present upstream of
ATG start codon with a possible signal peptide
present

40

ID-92

Clone RS-69 (partial sequence)

5 ATGAAAAAGAAAACATTCAGTGCTTATAACTTTTTAACGGCTCTTATCCTT
TGTCTTTTGACAGTGCTTTTATCTTTCCATTTTATTGGATTATGACAGGAG
CTTTTAA

MKKKTFSAYNFLTALILCLLTVLFIFPFYWIMTGAF

10

Sequence description:

15 A] Length: 110 bp -36 aa (Partial sequence)
B] Possible signal peptide with Shine Dalgarno
sequence directly upstream of the ATG start
codon.

20

ID-93

Clone RS-70

25

30 ATGACTGAGAACTGGTTACATACTAAAGATGGTTCAGATATTTATTATCGT
GTCGTTGGTCAAGGTCAACCGATTGTTTTTTTACATGGCAATAGCTTAAGT
AGTCGCTATTTTGATAAGCAAATAGCATATTTTCTAAGTATTACCAAGTT
ATTGTTATGGATAGTAGAGGGCATGGCAAAAGTCATGCAAAGCTAAATAC
CATTAGTTTCAGGCAAATAGCAGTTGACTTAAAGGATATCTTAGTTCATTT
AGAGATTGATAAAGTTATATTGGTAGGCCATAGCGATGGTGCTAATTTAGC
35 TTAGTTTTTCAAACGATGTTTCCAGATATGGTTAGAGGGGCTTTTGCTTAAT
TCAGGGAACCTGACTATTCATGGTCAGCGATGGTGGGATATTCTTTTAGTA
AGGATTGCCTATAAATTCCTTCACTATTTAGGGAAACTCTTCCGTATATG
AGGCAAAAAGCTCAAGTTATTTTCGCTTATGTTGGAGGATTTGAAGATTAGT
CCAGCTGATTTACAGCATGTGTCAACTCCTGTAATGGTTTTGGTTGGAAAT
AAGGACATAATTAAGTTAAATCATTCTAAGAACTTGCTTCTTATTTTCCA
40 AGGGGGGAGTTTTATTCTTTAGTTGGCTTTGGGCATCACATTATTAAGCAA
GATTCCCATGTTTTTAATATTATTGCAAAAAAGTTTATCAACGATACGTTG
AAAGGAGAAATTGTTGAAAAAGCTAATTGA

MTENWLHTKDGSDIYYRVVGQGQPIVFLHGNSLSSRYFDKQIAYFSKYYQVIV
MDSRGHGKSHAKLNTISFRQIAVDLKDILVHLEIDKVILVGHSDGANLALVFQ

TMFPDMVRGLLLNSGNLTIHGQRWWDILLVRIAYKFLHYLGKLFYPYMRQKA
QVISLMLEDLKISPADLQHVSTPVMVLVGNKDIKLNHKKLASYPFRGEFYSL
VGFGHHIIKQDSHVFNIIAKKFINDTLKGEIVEKAN*

5

Sequence description:

- 10 A] Length: 744 bp - 248 aa (full length gene)
 B] No obvious signal peptide, but Shine
 Dalgarno sequence upstream of the ATG start
 codon.

15

ID-94

20

Clone RS-71

25

ATGGTAGCAAAAGAGTTAGGTAAAAATAGCTTTACTATCCCAACTATTTGT
TCTAATTGCTCCGCAGGTACTGCCATTGCAGTTGTATATAATGATGACCAT
TCTTCTTAAGATACGGCTATCCCGAGTCTCCACTTCATATTTTTATCAATA
CACGGATCATTGCACAGGCACCAAGCAAATATTTTTGGGCTGGTATTGGGG
ACGGTATTTCAAAAGCCCCTGAAGTAGAACGTGCTACCTTAGAGGCTAAG
ACCAATAAACTACCACATACTGCAGTGTTAGGACAAGCAGTCGCTCTGTCT
TCAAAGGAAGCTTTTTATCAATTTGGTGAACAAGGTCTAAAAGACGTTGAA
GCTAATTTAGCTTCGCGTGCAGTTGAAGAAATTGCGCTTGATATCTTA

30

MVAKELGKNSFTIPTICSNCSAGTAIAVVYNDHDSFLRYGYPEsplHIFINTRIIA
QAPSKYFWAGIGDGISKAPEVERATLEAKTNKLPHTAVLGQAVALSskeAFY
QFGEQGLKDVEANLASRAVEEIALDIL

35

Sequence description:

- 40 A] Length: 405 bp - 135 aa (Partial sequence)
 B] No obvious Shine Dalgarno sequence upstream
 of the ATG start codon, probable signal
 peptide present at the N-terminus.

ID-95

Clone RS-73

5 TTGAGGGGAACTTACTGGAAAATTTCAAGCGATTGCGATAAAATAAATCTT
 GCAGAGTTTTCTAGAGAAAGGAGGTCAGATTTATTGGAGTGGCAAGATCT
 AGCGCAGTTACCTGTATCTATTTTTAAAGACTATGTTACAGATGCTCAAGA
 CGCGGAAAAACCTTTTATATGGACAGAAGTATTTTAAAGGGAGATTAATCG
 10 CTCAAATCAAGAAATTATTTTGCATATTTGGCCGATGACTAAGACAGTCAT
 TCTGGGGATGTTAGATCGAGAATTACCACATTTAGAATTAGCTAAAAAAG
 AAATCATCAGTCGTGGTTATGAACCAGTTGTTTCGGAATTTTGGAGGTCTCG
 CAGTTGTAGCTGATGAAGGAATTTTAAATTTTTCATTGGTTATTCCAGATGT
 TTTTGAGAGAAAATTGTCTATCTCAGATGGGTATCTTATAATGGTCGATTTT
 ATTAGAAGTATATTTTCGGATTTTATCAACCTATTGAGCACTTTGAAGTA
 15 GAGACCTCCTATTGTCCTGGTAAGTTTGATCTTAGTATAAATGGCAAAAAA
 TTTGCTGGCTTGGCTCAGCGCCGTATAAAGAATGGTATTGCGGTATCAATT
 TACCTTAGCGTTTGTGGCGATCAAAAAGGGCGGAGTCAAATGATTTTCAGAT
 TTTTATAAGATTGGTCTAGGTGATACGGGTAGTCCAATTGCTTATCCAAAT
 GTAGATCCTGAAATTATGGCTAATCTATCTGATCTATTAGATTGTCCTATG
 20 ACAGTAGAAGATGTTATTGATCGTATGTTGATTAGCCTTAAACAAGTAGGT
 TTTAATGATCGTTTACTGATGATTAGACCCGATTTAGTTGCAGAGTTTGAT
 AGATTTTCAGGCTAAGTCTATGGCTAATAAGGGGATGGTGAGCAGAGATGA
 ATAA
 25
 MRETYWKISSDCDKINLAEFSRERRSDLLEWQDLAQLPVSIFKDYVTDAQDAE
 KPFIWTEVFLREINRSNQEILHIWPMTKTVILGMLDRELPHLELAKKEIISRGYE
 PVVRNFGGLAVVADEGILNFSLVIPDVFERKLSISDGYLIMVDFIRSIFSDFYQPI
 30 EHFEVETSYCPGKFDLSINGKKFAGLAQRRIKNGIAVSIYLSVCGDQKGRSQMI
 SDFYKIGLGD TGSP IAYPNVDPEIMANLSDLLDCPMTVEDVIDRMLISLKQVGF
 NDRLLMIRPDLVAEFDRFQAKSMANKGMVSRDE*

Sequence description:

35 A] Length: 921 bp -307 aa (Full-length gene sequence)
 B] No obvious Shine Dalgarno sequence upstream
 of the TTG start codon or signal peptide
 visible. Actual start point may be a further
 40 85 bp downstream (TTG). This start point is
 preceded by a typical Shine-Dalgarno sequence.

ID-96

Clone RS-74

5

TTGGAAGGTTTACTTATTGCATTGATTCCCATGTTTGC GTGGGAAAGTATT
GGATTTGTTAGTAATAAAAATTGGAGGGCGTCCAAATCAACAAACATTTGG
AATGACTTTAGGAGCATTGCTATTTGCGATTATCGTATGGTTATTTAAACA
10 GCCAGAGATGACTGCCTCATTGTGGATTTTTGGTATCTTAGGTGGTATCCT
ATGGTCAGTCGGCCAAAATGGTCAATTTCAAGCAATGAAATATATGGGAG
TCTCTGTTGCTAATCCACTGTCAAGTGGTGCACAATTAGTAGGTGGAAGCC
TAGTTGGTGCTTTAGTCTTTTCATGAATGGACTAAGCCAATCCAATTTATTTT
AGGATTGACAGCGTTGACATTATTAGTTATCGGCTTCTATTTCTCAAGTAA
15 ACGTGATGTTTCAGAACAAAGCTTTGGCAACACATCAAGAGTTTTCAAAG
GATTTGCTACAATTGCTTATTCAACTGTAGGTACATCTCGTACGCAGTTTT
ATTTAACAAACATTATGAAGTTCGACGCTATGGCCGTCATTTTACCCATGGC
TGTTGGAATGTGTCTAGGTGCAATTTGTTTCATGAAGTTTCGTGTAACTTT
GAGGCTGTTGTTGTTAAAAATATGATTACAGGTCTCATGTGGGGCGTTGGT
20 AATGTCTTCATGTTATTGGCAGCAGCTAAAGCAGGGCTAGCAATTGCTTTT
AGTTTTTCTCAACTTGGAGTAATTATCTCTATTATTGGTGGTATTTTATTTT
AGGTGAGACAAAAACGAAGAAAGAGCAGAAATGGGTTGTCATGGGTATC
CTTTGTTTTGTTATGGGTGCTATATTACTTGGTATTGTTAAATCTTATTAA
25 MEGLLI ALIPMFAWESIGFVSNKIGGRPNQQTFGMTLGALLFAIIVWLFKQPEM
TASLWIFGILGGILWSVGQNGQFQAMKYMGVSVANPLSSGAQLVGGSLVGAL
VFHEWTKPIQFILGLTALTLLVIGFYFSSKRDVSEQALATHQEFSKGFATIAYST
VGYISYAVLFNNIMKFDAMAVILPMAVGMCLGAICFMKFRVNFEAVVVKNMI
TGLMWGVGNVFMLLAAAKAGLAIAFSFSQLGVIIISIGGILFLGETKTKKEQK
30 WVVMGILCFVMGAILLGIVKSY*

Sequence description:

35

A] Length: 867 bp - 289 aa (full-length gene)
B] Possible Shine Dalgarno sequence upstream of
GTG start codon, no obvious signal peptide
present.

40

ID-97

Clone RS-75

5 ATGACAACCTTACTACGAAGCTATAAACTGGAACGAAATTGAAGATGTTAT
 TGATAAATCAACTTGGGAAAAACTAACCGAACAATTTTGGCTCGATACAC
 GTATCCCTTTTATCAAATGACTTAGACGATTGGCGCAAACCTTTCCGCTCAAG
 AAAAAGATCTTGTTGGCAAGGTTTTTGGAGGCTTAACCCTACTTGATACCA
 TGCAATCAGAACTGGTGTGGAAGCTATTTCGTGCCGATGTTTCGCACGCCTC
 10 ACGAAGAAGCTGTCTTAAACAATATTCAATTCATGGAATCTGTTACGCTA
 AATCTTATTCTTCAATTTTCTCAACTTTAAATACTAAATCAGAAATTGAAG
 AAATTTTCGAGTGGACTAATAATAATGAGTTCCTTCAAGAAAAAGCACGT
 ATTATCAATGACATTTATGCTAATGGAAATGCCCTTCAAAAAAAGGTGGCT
 TCCACCTACCTCGAAACTTTCCTTTTTTATTCTGGCTTTTTTCACACCTCTTTA
 CTATTTGGGAAATAATAAGTTAGCAAATGTTGCTGAAATCATTAAATTAAT
 15 TATTCGTGATGAATCTGTACATGGTACTTATATCGGTTACAAATTCCAGCTT
 GGTTTTAACGAATTACCAGAAGATGAGCAAGAGAATTTTCGTGATTGGAT
 GTATGACCTCCTTTATCAGCTGTATGAAAACGAAGAAAAATACACCAAGA
 CACTTTATGATGGCGTAGGATGGACTGAAGAAGTTATGACCTTTTTACGCT
 ACAATGCTAATAAAGCTCTTATGAATTTAGGACAAGATCCTTTATTCCCAG
 20 ATACAGCAAATGATGTCAACCCAATTGTTATGAATGGTATTTCAACAGGAA
 CATCAAACCATGACTTCTTCTCTCAAGTAGGTAATGGTTACCTACTTGGTA
 GCGTTGAAGCTATGCATGATGATGACTATAACTATGGATTATAA

25 MTTYEAINWNEIEDVIDKSTWEKLTEQFWLDTRIPLSNDLDDWRKLSAQEK
 DLVGKVFGLTLLDTMQSETGVEAIRADV RTPHEEAVLNNIQFMESVHAKSY
 SSIFSTLNTKSEIEEIFEWTNNNEFLQEKARIINDIYANGNALQKKVASTYLETF
 LFYSGFFTPLYYLGNKLANVAEIIKLIIRDES VHGTYIGYKFQLGFNELPEDEQ
 ENFRDWMYDLLYQLYENEEKYTKTLYDGVGWTEEVMTFLRYNANKALMNL
 30 GQDPLFPDTANDVNPIVMNGISTGTSNHDFFSQVGNGYLLGSVEAMHDDYDYN
 YGL*

Sequence description:

35 A] Length: 960 bp - 320 aa (full length gene)
 B] Shine Dalgarno sequence present upstream of
 ATG start codon, but no signal peptide
 present.

40

ID-98

Clone RS-77 (partial sequence)

5 ATGAATTGGTCACGTATCTGGGAACTCGTAAAAATTAATATCCTTTATTCA
AACCCTCAGACTCTATCGGCACTAAGAAAAAAGCAAGAAAAGCATCCTAA
AAAAGAATTTTCAGCTTATAAATCCATGTTTAGAAATCAGTTATTTTCAGAT
TTTGCTCTTTTCAATAATTTATGTATTTCTCTTTGTATCACTTGATTTTAAAG
AATATCCGGGCTATTTACGTTCTACATTGGTATCTTTACACTAGTATCCAT
10 TATCTACTCTTTTATTGCGATGTACAGTGTTTTCTATGAGAGTGACGATGTT
AA

MNWSRIWELVKINILYSNPQTLSALRKKQEKHPKKEFSAYKSMFRNQLFQILL
FSIIYVFLFVSLDFKEYPGYFTFYIGIFTLVSIISFIAMYSVVFYESDDV

15

Sequence description:

- 20 A] Length: 311 bp - 103 aa (Partial sequence)
B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide at
N-terminus.

ID-99

25

Clone RS-78 (partial sequence)

30 TAATCTTTTAGTCAACGGAGCAACAGGAAAATTGCAGGCTATGCGACAGA
TATTCCACCACATAATTTAGCAGAAGTCATTGATGCTGTCGTGTACATGAT
TGATCACCCCTAAAGCTAAATTAGATAAATTAATGGAATTTCTACCTGGTCC
AGATTTTCCAACCTGGCGCTATCATTCAAGGAAAAGATGAAATTCGTAAGG
CATATGAGACTGGTAAGGGGAGAGTAGCGGTTTCGCTCGCGAACTGCTATT
35 GAAACCTTAAAAGGTGGTAAGAAACAAATTATTGTTACTGAAATTCCTTAT
GAAGTTAAT

SFSQRSNRKIAGYATDIPPHNLAEVIDAVVYMIDHPKAKLDKLMEFLPGPDFPT
GAIQGGKDEIRKAYETGKGRVAVRSRTAIETLKGGKKQIIVTEIPYEVN

40

Sequence description:

- A] Length: 312 bp - 104 aa (Partial sequence)
B] No obvious Shine Dalgarno sequence or a

signal peptide. Both N- and C- termini of ORF
yet to be elucidated.

5 ID-100

Clone RS-79

10

ATGGGACGTAAGTGGGCCAATATTGTTGCCAAAAAGACTGCTAAAGATGG
TGCTAACTCAAAAGTATACGCTAAATTCGGTGTGAAATATATGTTGCTGC
AAAGCAAGGTGAACCAGACCCCGAGTCAAACCTCAGCTCTAAAATTCGTTT
TGGACCGTGCTAAGCAAGCACAAAGTTCCAAAGCATGTTATTGATAAAGCG
15 ATTGATAAAGCCAAAGGAAACACAGATGAAACTTTTCGTAGAGGGACGCTA
TGAAGGTTTTTGGTCCAAATGGTTCAATGATTATTGTGGATACTTTGACATC
AAATGTTAACCGTACGGCAGCAAATGTACGTACTGCTTACGGTAAGAACG
GTGGCAATATGGGAGCTTCAGGATCGGTATCCTACTTATTTGATAAAAAAG
GTGTCATCGTTTTTGTCTGGTGATGATGCTGACACTGTCTTCGAACAATTACT
20 TGAAGCGGATGTAGACGTAGATGATGTTGAAGCAGAAGAGGGAACAATA
ACAGTTTATACCGCCCCAACAGATCTTCATAAAGGTATCCAAGCACTTCGC
GATAATGGTGTAGAAGAATTCCAAGTTACTGAACTTGAAATGATTCCTCAA
TCAGAAGTAGTATTGGAAGGTGATGACCTTGAAACTTTTGAAAAGCTT

25

MGRKWANIVAKKTAKDGANSKVYAKFGVEIYVAAKQGEPPDPESNSALKFVL
DRAKQAQVPKHVIDKAIDKAKGNTDETFVEGRYEGFGPNGSMIIVDTLTSNV
NRTAANVRTAYGKNNGNMGASGSVSYLFDKKGIVVFAGDDADTVFEQLLEA
DVDVDDVEAEEGTITVYTAPDLHKGIALRDNGVEEFQVTELEMIPQSEVVL
EGDDLETFEKL

30

Sequence description:

- 35 A] Length: 654 bp - 218 aa (Partial sequence)
B] Possible Shine Dalgarno sequence upstream
of ATG start, no obvious signal peptide

40 ID-101

Clone RS-80

TTGGAGAAATATTTGAAGAACCCGATTACATGGATTGGATTAGTTCTTGTG
GTTACGTGGTTTTTAACTAAAAGTAGTGAATTTTGGATTGTTGGTGTGTGTG
TCTTGTTGTTAGTATTTGCTAGTCAAAGTGAT

5

MEKYLKNPITWIGLVLVVTWFLTKSSEFLIFGVCVLLLVFASQSD

Sequence description:

10

A] Length: 135 bp - 45 aa (partial sequence)
B] Shine Dalgarno sequence upstream of TTG
start codon with possible signal peptide
evident at N-terminus.

15

ID-102

20

Clone RS-81

ATGACACAATCAGATGCATATCTCTCGTTGAACGCGAAGACACGCTTTAGA
GATCGCACAGGTAATTATCATTTTACTTCGGATAAAGAGGCTGTTGAACAA
TATATGATAGAACATGTTGAACCTAATACGATGGTGTTCACATCACTAATT
GAAAAGCTAGATTATTTGGTTTCTAATAACTACTATGAATCGGACCTTCTA
AAACAATATAACCTTGAGTTTATTTGCCAAATTTTGGAGCATGCATACGCT
AAGAAATTTGCTTTTCTAAATTTTATGGGGGCTTTAAAATTTTATAATGCTT
ATGCTCTTAAT

25

30

MTQSDAYLSLNAKTRFRDRTGNYHFTSDKEAVEQYMIEHVEPNTMVFTSLIE
KLDYLVSNYYESDLLKQYNLEFICQIFEHAYAKKFAFLNFMGALKFYNAYA
LN

35

Sequence description:

40

A] Length: 318 bp - 106 aa (Partial sequence)
B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide

ID-103

Clone 2-11A

5

ATGGTATTTATGGCAAATAAGAAAAAACAAGGAAAGAAAACCAGAA
GACCTACTAAGGCAGAAATAGAGCGTCAAAGAGCTATTCAAAGGATGATT
ACTGCTCTTGTTTTAACAAATTATTCTCTTCTTTGGTATTATCAGATTAGGTA
10 TTTTGGTATTACAGTCTATAACGTCATCCGTTTTATGGTAGGTAGCTTGGC
TTACTTATTTATTGCGGCAACTTTAATCTACCTTTATTTCTTTAAATGGTTG
CGAAAGAAAGATAGCTTAGTAGCAGGTTTTTTGATAGCTTCTTTAGGATTA
TTGATTGAGTGGCATGCTTACCTTTTCTCAATGCCTATTTTGAAAGATAAA
GAAATTTTGC GTTCAACTGCTCGATTAATTGTGTCTGATTTAATGCAATTTA
15 AAATCACTGTTTTTGCCGGTGGAGGTATGTTGGGTGCTTTGATTTACAAGC
CAATTGCTTTTCTCTTTTCTAATATTGGTGCCTATATGATTGGTGTCTCTTC
ATCATTTTGGGTCTCTTTTAAATGAGTTCTCTGGAAGTTTATGACATCGTCG
AATTTATTAGAGCTTTTAAAAATAAAGTGGCAGAGAAGCACGAGCAAAAT
AAAAAGGAGCGTTTTGCTAAGCGAGAGATGAAAAAAGCAATCGCTGAACA
20 AGAGCGCATAGAGCGTCAAAAAGCTGAAGAAGAAGCTTATTTAGCTTCGG
TTAATGTAGACCCTGAAACGGGTGAGATTCTAGAGGATCAAGCTGAGGAC
AATTTGGATGATGCGCTACCACCTGAGGTAAGTGAAACATCAACTCCGGT
ATTTGAGCCAGAGATCCTTGCTTATGAGACATCGCCTCAAATGATCCTTT
ACCAGTAGAGCCGACAATTTATTTAGAAGACTATGATTCGCCGATTTCCTAA
25 TATGAGAGAAAATGATGAGGAAATGGTTTATGATTTAGATGATGATGTAG
ATGATAGTGATATAGAAAATGTCGACTTTACACCTAAAACGACACTGGTTT
ATAAATTACCAACGATAGATTTATTTGCACCAGATAAGCCTAAAAATCAAT
CCAAAGAAAAGGATTTAGTCCGAAAGAATATCAGAGTTTTAGAAGAAACA
TTTAGAAGTTTTGGTATCGATGTAAAAGTAGAACGTGCTGAAATTGGACCA
30 TCAGTTACTAAATATGAAATTAAACCAGCAGTTGGAGTTCGTGTGAATCGT
ATTTCAAATCTATCTGACGACCTAGCTCTTGCTCTTGACAGCAAAAGATGTG
CGTATAGAAGCACCAATTCCTGGAAAATCATTAAATAGGTATTGAAGTTCCT
AACTCAGAAATTGCAACGGTTTCTTTCCGCGAACTTTGGGAACAATCTGAT
GCCAATCCTGAAAACCTTTTAGAAGTACCACTAGGAAAAGCTGTTAACGG
35 CAATGCTCGCAGTTTTAACTTAGCTAGAATGCCGCATCTTTTGGTAGCTGG
TTCAACTGGTTCAGGTAAATCTGTGGCAGTTAATGGAATTATTTCAAGTAT
TTTGATGAAGGCACGTCCAGATCAAGTTAAGTTTATGATGATTGATCCCAA
AATGGTTGAATTATCTGTTTATAATGATATTCCACATTTATTAATCCCTGTT
GTAACCAATCCGCGTAAAGCAAGTAAGGCACTCCAAAAAGTTGTTGATGA
40 AATGGAAAATCGATACGAGTTATTTAGCAAAATTGGTGTGCGTAATATAG
CAGGTTATAATACAAAGGTTGAAGAGTTTAATGCTTCCTCTGAGCAAAAC
AAATGCCTTTGCCTTTAATCGTTGTCAATTGTAGATGAATTGGCTGACTTGAT
GATGGTTGCTAGTAAAGAAGTTGAAGATGCTATTATTCGTTTGGGGCAAAA
AGCACGTGCTGCAGGTATCCATATGATTCTTGCAACTCAACGTCCATCCGT

AGATGTTATTTCTGGTTTGATTAAAGCAAATGTTCCGTCGCGTATTGCATTT
GCTGTTTCAAGTGGTACTGATAGCCGTACGATCCTTGATGAAAATGGTGCT
GAAAAGCTCTTGGGACGGGGTGACATGCTCTTTAAGCCTATTGATGAGAAT
CATCCAGTACGACTACAAGGTTCCCTTTATTTTCAGATGATGATGTTGAAAGG
5 ATCGTTGGTTTTATCAAAGACCAAGCCGAGGCTGACTATGATGATGCCTTT
GATCCTGGAGAAGTATCTGAAACAGATAACGGCTCTGGTGGTGGCGGCGG
AGTACCTGAAAGTGATCCTCTTTTTGAAGAAGCCAAGGGACTCGTTTTAGA
GACGCAAAAAGCAAGTGCCTCAATGATTCAACGCCGATTGTCTGTTGGTTT
CAATAGAGCAACAAGACTAATGGAAGAATTAGAAGCAGCGGGGGTTATTG
10 GTCCAGCAGAAGGAACCAAGCCACGAAAAGTTTTAATGACTCCAACCTCCG
AGTGAATAA

MVFMANKKKTKGKKTRRPTKAEIERQRAIQRMITALVLTIIFFGIIRLGIFGIT
VYNVIRFMVGSLAYLFIAATLIYLYFFKWLRKKDSLAVAGFLIASLGLLIEWHA
15 YLFSMPILKDKEILRSTARLIVSDLMQFKITVFAGGGMLGALIYKPIAFLFSNIG
AYMIGVLFILGLFLMSSLEVYDIVEFIRAFKNKVAEKHEQNKKERFAKREMK
KAIAEQERIERQKAEEEEAYLASVNVDPETGEILEDQAEDNLDDALPPEVSETST
PVFEPEILAYETSPQNDPLPVEPTIYLEDYDSPINMRENDEEMVYDLDDVDVDD
SDIENVDFTPKTTLVYKLPTIDLFAPDKPKNQSKEKDLVRKNIRVLEETFRSFGI
20 DVKVERAEIGPSVTKYEIKPAVGVRVNRISNLSDDLALALAAKDVRIEAPIPGK
SLIGIEVPNSEIATVSFRELWEQSDANPENLLEVPLGKAVNGNARSFNLARMPH
LLVAGSTGSGKSVAVNGISSLMKARPDQVKFMMIDPKMVELSVYNDIPHLLI
PVVTNPRKASKALQKVVDENENRYELFSKIGVRNIAGYNTKVEEFNASSEQK
QMPLPLIVVIVDELADLMMVASKEVEDAIIRLGQKARAAGIHMILATQRPSVD
25 VISGLIKANVPSRIAFAVSSGTDSTRITLDENGAEKLLGRGDMLFKPIDENHPVRL
QGSFISDDDVERIVGFIKDQAEADYDDAFDPGEVSETDNGSGGGGGVPESDPL
FEEAKGLVLETQKASASMIQRRLSVGFNRATRLMEELEAAGVIGPAEGTKPRK
VLMTPTPSE*

30

Sequence description:

35 A] Length: 2451 bp - 817 aa (Full-length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

40 ID-104

Clone 2-18/22b

ATGTCACAAGAGCAAGGAAAAATTTATATTGTAGAAGATGATATGACGAT
TGTGTCACCTTTTAAAAGATCATTTATCAGCTAGCTATCATGTCTCTAGTGTC
AGCAATTTTCGTGATGTGAAACAAGAAATTATCGCATTTC AACCCGATTG
5 AACTAATGGATATTACGTTACCCTATTTTAATGGTTTTTACTGGACTGCAG
AATTGCGTAAGTTTTTAACAATTCCTATTATTTTCATTTTCATCTAGTAATGA
TGAAATGGATATGGTTATGGCATTAAATATGGGGGGTGATGACTTTATTTC
AAAACCATTCTCTCTAGCTGTATTAGATGCTAAGCTAACTGCTATTTTAAAG
GAGAAGTCAACAATTTATCCAACAGGAATTAAGCTTTTGGGGGATTACGTT
10 GACAAGAGAAGGGTTATTGTCTAGCCAAGATAAAGAGGTTATTTTATCGC
CAACAGAAAATAAAATCCTATCTATCTTGCTCATGCATCCTAAACAAGTAG
TCTCAAAAGAGTCTCTATTAGAGAACTTTGGGAAAATGATAGTTTTATTG
ATCAAAATACACTTAATGTTAATATGACACGCTTACGTAAAAAAATTGTCC
CAATAGGTTTTGATTACATTCATACAGTGAGAGGAGTTGGGTATTTACTAC
AATGA

15

MSQEQGKIYIVEDDMTIVSLLKDHLASYSYHVSSVSNFRDVKQEIIAFQPDILM
DITLPYFNGFYWTAELRKFLTIPFISSNDEMMDMVMALNMGGDDFISKPFSLA
VLDAKLTAILRRSQQFIQQELTFGGFTLTREGLLSSQDKEVILSPTENKILSILLM
20 HPKQVVSKESSLLEKLWENDSFIDQNTLVNMTRLRKKIVPIGFDYIHTVRGVG
YLLQ*

25 Sequence description:

A] Length: 669 bp - 223 aa (full-length gene
sequence)

30 B] Shine Dalgarno sequence present upstream of a GTG start codon.
Was not identified directly by LEEP. This gene was found upstream of
gene ID-10 described in WO 00/06736.

ID-105

35

Clone 2-20

ATGTATCAAACCTCAGACAAATAAGGAAAAATTTGTTTTATTTTTGAAATTA
TTTATCCCAGTATTGATTTATCAATTTGCTAATTTTTCAGCTACTTTTATTGA
40 TTCGGTTATGACTGGACAGTATAGTCAGCTACATTTGGCAGGTGTGTCAAC
TGCTAGTAATTTATGGACTCCGTTTTTCGCTTTATTAGTAGGTATGATTTCA
GCATTAGTACCAGTAGTTGGTCAACATTTGGGTAGAGGAAATAAAGAACA
AATTCGCACAGAATTTTCATCAATTTCTATATTTAGGTTTGATACTGTCCTTA
ATATTATTTTAAATCATGCAATTTATTGCTCAACCTGTCTTGGGGAGTTTGG

GTTTAGAAGATGAAGTTCTAGCAGTTGGTCGTGGTTATTTAAATTATATGT
 TGATTGGAATCATGCCGCTGGTGTGTTTAGCATTTGCCGTTCAATCCTTTGA
 TGCATTGGGGTTAACAAGGTTATCTATGTATCTGATGCTTTTAATTCTACCC
 5 TTTAATTCATTTTTTAATTATATGCTTATCTACGGTAAATTTGGGTATGCCTA
 GACTAGGAGGTGCGGGGGCAGGTCTTGGAACCTTCTTTAACTTATTGGGGCTA
 TTTTTATTGGTATTATTATTGTGATGTCACCTTCATCCTCAAATTAACATA
 TCATATATGGACTCTGGAAAGAATAAAAGCTCCTTTGATTATTGAAGATAT
 TCGATTGGGATTACCGATTGGTTTACAAATTTTTGCAGAAGTTGCAATTTTT
 GCAGTAGTAGGCTTATTCATGGCAAATTTTCTTCAATCATTATTGCAGCA
 10 CATCAGGCTGCTATGAATTTTTCATCATTAAATGTATGCATTTCTTTAAGTA
 TTTCCACTGCTCTAGCTATTACAATATCGTTTGAAGTAGGGGCAGAGCGCT
 TTCAGGACGCAACCACTTATAGTAGGATAGGACGCTTAACAGCGGTAGGG
 ATTACATCAGGAACCTTACTATTTTTATTTCTATTTTCGTGAGAATGTAGCAG
 CAATGTATAATAGTGCCCTCACTTTGTGCGCTATTACAGCTCAATTCCTAAC
 15 TTATAGTCTCTTTTTCCAGTTTGCAGATGCTTATGCAGCTCCTGTACAGGGG
 ATTTTACGAGGCTATAAGGATAACAACAAACCATTATGATCGGTGCGGG
 CTCTTATTGGTTATGTGCTTTGCCATTAGCGGTTATCTTAGAAAAAATAG
 CCAGTTAGGTCCGTTTGCCTATTGGATTGGTTTAATCACAGGTATTTTTGTT
 TGTGGTCTATTTCTAAACCAACGTCTGCAAAAGATTAAGAAGTTGTATTAT
 20 TAA

MYQTQTNKEKFVLFLKLFIPVLIYQFANFSATFIDSVMTGQYSQLHLAGVSTAS
 NLWTPFFALLVGMISALVPVVGQHLGRGNKEQIRTEFHQFLYLGLILSLILFLI
 25 MQFIAQPVLGSLGLEDEVLAVGRGYLNYMLIGIMPLVLFSICRSFFDALGLTRL
 SMYLMILLIPNSFFNYMLIYGKFGMPRLGGAGAGLGTSLTYWAIFIGIIIVMS
 LHPQIKTYHIWTLERIKAPLIIEDIRLGLPIGLQIFA EVAIFAVVGLFMKFSSIIIA
 AHQAAMNFSSLMYAFPLSISTALAITISFEVGAERFQDATTYSRIGRLTAVGITS
 GTLLFLFLFRENVAAAMYN SAPHFVAITAQFLTYSLFFQFADAYAAPVQGILRG
 YKDTTKPFMIGAGSYWLCALPLAVILEKNSQLGPFAYWIGLITGIFVCGLFLNQ
 30 RLQKIKKLYY*

Sequence description:

- 35
- A] Length: 1341 bp - 447 aa (full length gene)
 - B] Shine-Dalgarno sequence present upstream of
 - ATG start codon, There is a potential signal
 - peptide sequence
- 40

ID-106

Clone 2-4A

5 TTGCTAGTTTCTTCTCTAGTTTCTTGTTCAATTTTTCTTGTCATTTTCGTCGTT
GTCTTCATCAACACGAAATAAGTCTATAAACTTATCAAATAATTTTCATAGA
CTTATTATATCAATTTTCAATAAAATGCTATAATAAAACCATGTCATTTTCA
TTAAAAATTAGAAATCCATACGGTGAACATACCGTTAAAGAACTCCTTGA
10 AGATTATTTTTTGGATTCCACGTAAGATTAGACATTTTTTGCCTGTTAAAAAA
CATGTACTTATAAACAATGAATTCATTAATTGGCAAACCTGTCGTCCAAGAA
AACGATACTATTACCTTAATCTTTGATGATGAGGATTACCCTACTAAAAAA
ATTCCTCTGGGCAGAGCAGAGCTTATTGATTGTCTTTATGAGGATGAACAT
CTTATTATCGTTAATAAACCTGAAGGTATGAAAACCTCACGGTAACCAACCA
AATGAAATAGCACTGTTAAATCATGTATCTGCCTATTCTGGACAAACATGC
15 TATGTTGTTTCATCGCCTAGATATGGAGACCAGTGGAGCTGTTTTATTTGCT
AAAAATCCATTTATACTTCCCCTTATCAATCAACGCTTAGAACGAAAAGAA
ATTTGGCGTGAATATTGGGCTTTAGTTGAAGGAAAATTTTCACCTAAGCAT
CAAGTTTTTGAGAGACAAAATTGGACGGAACCGTCATGACAGACGTAACG
AATCATTGATTCTAAAAACGGTCAACATGCTATGACAATCATTGACGTTTT
20 GAAGTATATCCAAAATAGTAGTCTCATAAAATGCCGACTGGAAACCGGAA
GAACCCATCAAATTCGCATTCACTTATCTCATCACGGACATCCTTTAATAG
GAGATCCCCTCTACAACCCTTCTTCTAATAATGAAAGGTTAATGCTACACG
CTCACCGATTGACTCTATCCCATCCATTAACCTTGCGAACTATTAGCGTAG
AGGCCCTTCATCTACTTTTCGAGAAGGTTTTAAACAATTATAAAAAAGGAG
25 TTGGATAA

MLVSSLVSCSFFLVISSLSSSTRNKSINLSNNFIDLLYQFSIKCYNKTMSFSLKIR
NPYGEHTVKELLEDYFLIPRKIRHFLRVKKHVLINNEFINWQTVVQENDTITLIF
DDEDYPTKKIPLGRAELIDCLYEDEHLIIVNKPEGMKTHGNQPNEIALLNHVSA
30 YSGQTCYVVHRLDMETSGAVLFAKNPFILPLINQRLERKEIWREYWALVEGKF
SPKHQVLRDKIGRNRHRRKRIIDSKNGQHAMTIIDVLKYIQNSSLIKCRLETG
RTHQIRIHLSHHGHPLIGDPLYNPSSNNERLMLHAHRLTLSHPLTCETISVEAPS
STFEKVLNNYKKGVG*

35

Sequence description:

40

A] Length: 1029 bp - 343 aa (Full length gene sequence)
B] No obvious Shine-Dalgarno sequence upstream
of the putative TTG start codon. Possesses a
potential leader peptide sequence.

ID-107

Clone 2-54

5

GAAGTAAATGCAACTCAACCTAATAATAGAACTACCTATATTATACCCGAA
AGCAGTCATTCCATTGCAGAACACAGAGATTCCTGATAGAATCAAAGGG
TTCTTCGGTTGCATTACTTAATAGCGATGAATTTAGAAAGACAGCGGGAGA
10 GGATAGAGGTTTTGAAAGGGATAAGTTGAGGTCTTTGGATATCATTCTAA
GGGAGATTTATCGACAAGTAATGTCATAGGTAATACGGACATTGCTAGTC
AGATATCGTTGGGCTTTAAAAAGAATGCGATGCAGGAACACCATCTTACT
AAAACATTCTCTCAAAAGGATGGAAAGTTATCGTCTGTTATAGAGGGGAT
GCTTGCTATTGGCAAAGAGAAAGTAGAGAAAGAAATAAAATATAGTGGTA
15 ATTTATGGCAAAAATTTAAAGCTAAGGCACACTGCCTTGTTTGCTGTGTTG
ATAATTTGAATTTTGAAGATATAAAATCTTATTTTCAATATTATTGTCATCT
AAACCATCAGCTCAAATTACCTAAAGGTGCTATACTTTCTGCTAAAACAGA
AGTATATAGGGGAGGAGATTTTGGGAGAAAAAATAAAGATAATGTGTTTG
GTTACCGTATCCCCTCATTATTGAAAACCCAAAAAGGAACCTTACTTGCGG
20 GAGCTGATGAAAGAATTGAGCAAGCTTGTGATTGGGGAAACATAGGAATG
GTTATTCGCCGTAGTGAGGATGATGGTGTCACTTGGGGAAAAAGAGAAAC
TATTGTCAATCTCCGTAATAACCCTAGAGTTCGCTAGTTACTAGTGGTGA
CTATAGTGGCTCACCTATTAATATGGATATGGCATTAGTTCAAGATACTAG
CTCCAAGACGAAACGTATTTTTTCAATATATGATATGTTTCCAGAAGGAAG
25 AGGCGTTATTAGTATTGCTAACACACCTGAAAAAGAATATACCCAAATCG
GAGGACAGTCTTATCTTAATTTATATAATAATGGAAAGAAATCGAAGGTTT
TACTATCCGTGACAAAGGTATTGTATATAATTTTAAAGGGAAAAAGACTG
ATTATCATGTTATAACAGAACTACTAAAAGTGACCATTCAAATCTAGGGG
ATATTTATAAGGGAAAACAGCTACTTGGAATATATATTTTACAAAACATA
30 AAACGTCACCATTTTCGTTTAGCAAAATCAAGCTATGTGTGGATGTCATATA
GCGATGATGATGGTAGGACATGGTCATCACCTAGAGATATAACAGCAAGT
CTTCGTCAGAAAGGCATGAAATTTTTGGGAATAGGACCTGGAAAAGGTAT
AGTTTTAAATGGGGGCCACACGCTGGTCGTATTATTATTCCTGCCTATTCT
ACGAATTGGAAATCTCATCTAAGAGGTTCACAATCTTCACGCCTAATTTAT
35 TCAGACGACCATGGAAAAACGTGGCATACTGGAAAAGCAGTTAATGATAA
CCGTATACTTTCTAATGGTGAAAAAATTCACTCCTTAACAATGGATAATAA
AAAAGAACAAAATACAGAATCCGTACCCGTTCAATTGAAAAATGGGGACA
TTAAGTTATTTATGAGGAATCTAACTGGTAACCTAGAAGTAGCCACAAGTA
AAGACGGCGGGGAGACTTGGCAAAACCATGTTAAACGATATAAGGAAATT
40 CATGATGCTTACGTCCAACCTATCAGCTATTCGCTTTGAGCATGACAAAAA
GAGTATATTTTATTAGTGAATGCTAATGGGCCAGGGAAGAAGTGCCAAGA
TGGATATGCACGTCTAGCGCAAGTTAATCGAAATGGTAGTTTTAAGTGGTT
ATATCACCATCACATTCAAGATGGTTCGTTTGCTTACAACCTCTGTTCAACA
ACTTAATAATGATCAATTTGGTGTCTTTATGAACATAGAGAAAAACATCA

AAATAGTTTTACTTTAAATTACAAAGTTTTTAAATTGGAGTTTTCTTAGTCAA
AATACAGAGAAGCAAGGCACTTTATGGGAGAAAATGGCAGCAAATTGGCA
TGTTTTGTTTAAATTTTATTTATGA

5 ELNATQPNNRTTYIIPESHSIAEQQRFLIESKGSSVALLNSDEFKRTAGEDRGF
ERDKLRSLDIIPKGDLSTSNVIGNTDIASQISLGFKKNAMQEHLTKTFSQKDG
KLSSVIEGMLAIGKEKVEKEIKYSGNLWQKLKAKAHCLVCCVDNLNFEDIKS
10 YFQYYCHLNHQLKLPKGAILSaktevyrGGDFGRKNKDNVFGYRIPSLLKTQ
KGTLLAGADERIEQACDWGNIGMVIRSEDDGVTWGKRETIVNLRNNPRVPL
VTSGDYSGSPINMDMALVQDTSSKTKRIFSIYDMFPEGRGVISIANTPEKEYTQI
GGQSYLNLYNNGKSKSVFTIRDKGIVYNFKGKKTDYHVITETTKSDHSNLGDI
YKKGQLLGNIYFTKHKTSPFRLAKSSYVWMSYSDDDGRTWSSPRDITASLRQ
15 KGMKFLGIGPGKGIVLKWGPHAGRIIPAYSTNWKSHLRGSQSSRLIYSDDHG
KTWHTGKAVNDNRILSNGEKIHSLTMDNKKEQNTESVPVQLKNGDIKLFMRN
LTGNLEVATSKDGGGETWQNHVKRYKEIHDAYVQLSAIRFEHDKKEYILLVNA
NGPGKKCQDGYARLAQVNRNGSFKWLYHHHIQDGSFAYNSVQQLNNDQFG
VLYEHREKHQNSFTLNYKVFNWSFLSQNTEKQGTLEWKMAANWHVLFKFYL
*

20 Sequence description:

- A] Length: 2052 bp - 684 aa (partial gene sequence)
- B] N-terminus has yet to be determined

25 ID-108

30 Clone 2-61

ATGCCTAAATTAATCGTATCTTTCCTCTGCATTTTATTATCCCTGACTTGTG
TAAACTCTGTGCAAGCTGAAGAACATAAAGATATTATGCAAATTACCCGA
35 GAAGCCGGATATGATGTTAAAGATATTAATAAACCTAAAGCGTCTATCGTT
ATTGACAATAAAGGTCATATTTTGTGGGAAGATAACGCCGATTTAGAACGT
GATCCCGCTAGCATGTCTAAAATGTTTACTTTATATTTACTATTTGAAGACT
TAGCTAAAGGAAAAACAAACCTCAACACCACAGTGACTGCAACAGAAACA
GACCAAGCCATAAGTAAGATTTATGAAATTAGTAATAACAATATTCATGCT
40 GGGGTTGCTTATCCTATTCGTGAAGTGAATGACTATGACGGCTGTCCCGTCA
TCTAATGTAGCAACTATTATGATTGCTAACCCTTATCACAAAACAATCCT
GACGCCTTTATTAAACGAATCAATGAAACCGCCAAGAACTCGGTATGAC
AAAACTCACTTTTATAACCCCAAGTGGGGCGGTAGCGAGTGCTTTTAAATGG
ACTTTACTCCCCAAAAGAATACGATAACAATGCTACTAACGTTACGACTGC

5

15

25

30

35

Clone 45

40

GGCTTTAGCAAATGGTAAGCACGTTCTTTGCGAAAAATCTATTACTTTAAA
 TAGTACTGAGCTTAAAGAAGCCATAGATTTAGCCGAAACTAACCATGTTGT
 CTTAGCTGAAGCCATGACTATTTTTCATATGCCAATTTACCGCCAATTTAAA
 AACATTAGTTGATAGTGGAAAATTAGGACCGTTAAAAATGATTCAAATGA
 5 ATTTCCGGAAGTTATAAAGAATATGATATGACTAACCGTTTTTTCAGTCGTG
 ACCTAGCAGGCGGTGCTTTGCTGGACATTGGTGTGTTATGCACTTTCTTGTAT
 TCGCTGGTTTATGTCAGAAGCACCTCACAACATTACCTCTCAAGTTACATT
 TGCACCAACAGGGGTTGATGAACAAGTTGGTATCCTACTAACCAACCCAG
 CAAATGAGATGGCGACTGTCAGCCTTAGTTTACATGCAAAACAACCTAAA
 10 CGAGCAACTATCGCTTACGATAAAGGCTACATTGAACTTTTTGAATATCCG
 CGAGGACAAAAGGCAGTTATTACTTATACTGAGGATGGGCATCAAGATAT
 TATCGAAGCTGGCAAAACTGAAAATGCTCTCCAATATGAGGTAGCTGATA
 TGGAAGAAGCCATTTTCAGGAAAACTAACCACATGTACTTAAACTATAACC
 AAAGATGTTATGGATATCATGACACAGCTACGTCAAGAATGGGGATTAC
 15 CTACCCAGAAGAAGAAAAATGA

MTEKYYNWATLGTGVIANELAQALEARQKLYSVANRTYDKGLEFANKYGI
 QKVYDHIDQVFEDPEVDIIYISTPHNTHISFLRKALANGKHVLCESITLNSTEL
 20 KEAIDLAETNHVVLAEAMTIFHMPYRQLKTLVDSGKLGPLKMIQMNFSGSYK
 EYDMTNRRFFSRDLAGGALLDIGVYALSCIRWFMSEAPHNITSQVTFAPTGVDE
 QVGILLTNPANEMATVSLSLHAKQPKRATIAYDKGYIELFEYPRGQKAVITYT
 EDGHQDIIEAGKTENALQYEVADMEEAISGKTNHMYLNYTKDVMMDIMTQLR
 QEWGFTYPEEEK*

25 Sequence description:

30 A] Length: 984 bp - 328 aa (full length gene)
 B] Shine Dalgarno sequence present upstream of
 ATG start codon, possesses a potential signal
 peptide

35 ID-110

Clone 2-2

40 GTGTATTCTCCTGTAAATCTTCTAAAGGAAAAGTGATATTGTAAAAAGT
 GATTTTCTAAAGAGCTTCATAGAAAGGAGAGGAAATATTTGTTTT
 MYSPVKSSKGKVILLKSDFLKSFIERRGNICF

Sequence description:

- 5 A] Length: 96 bp - 32 aa (partial sequence)
 B] GTG start codon - no obvious Shine-Dalgarno
 sequence
 Possesses a potential signal peptide

10

ID-111

15 Clone 2-3

20 AAATACTGTATCATTGCAACCTCAAATGCAGGTTTTGGAAACGAAGCATTT
 ACAGGTGACAGCGATAAAGACTTGAAAATTATGGAACGAATTTCTCCATA
 TTTCCGTCCAGAATTTCTAAATCGTTTCAATGGTGTTATTGAATTCTCTCAC
 CTAAGCAAAGATGACTTAAGCGAAATTGTAGATTTGATGCTTGATGAAGTT
 AACCAAACAATTGGCAAAAAAGGAATTGACCTTGTGGTAGATGAAAATGT
 TAAATCACACTTAATTGAACTGGGTTATGACGAAGCAATGGGAGTACGTC
 CATTGCGCCGTGTCATCGAGCAAGAAATTCGAGATCGCATCACAGACTACT
25 ATCTCGATCATAACAGACGTTAAACACCTAAAAGCTAATTTGCAAGATGGCC
 AAATCGTCATTTCTGAAAGATAA

30 KYCHATSNAFGNEAFTGDSKDLKIMERISPYFRPEFLNRFNGVIEFSHLKSD
 DLSEIVDLMLDEVNQTIGKKGIDLVVDENVKSHLIELGYDEAMGVRPLRRVIE
 QEIRDRTDYLDHTDVKHLKANLQDGQIVISER*

Sequence description:

- 35 A] Length: 429 bp - 143 aa (partial sequence)
 B] N-terminus yet to be elucidated. This gene
 was not in frame with nuc

40

ID-112

Clone 2-5

5 ATGTCAATGAATTTTTCATTTTACCACAATATTGGTCCTATTTTAATTATG
 GTGTGATGGTAACCATATGATTTCAACATGTGTTGTTTTTTTGGAACTAT
 TATAGGCGTGTTAATTGCTTTAGTAAAGCGTACTAATTTACATTTTCTCACA
 ATATTAGCTAATTTCTATGTATGGGTATTTTCGTGGGACACCGATGGTAGTT
 10 CAAATTATGATTGCTTTTCGCATGGATGCATTTTAACAATTTACCAACAATT
 AGCTTTGGTGTTTTAGATTTAGATTTTACACGACTTTTACCTGGTATCATT
 TCATTTCCTTAAATAGTGGTGCCTATATTTTCGGAAATTGTACGTGCAGGGA
 TTGAGGCTGTACCATCTGGACAAATAGAAGCAGCTTACTCGTTGGGGATTC
 GACCTAAAAATACACTTCGCTATGTTATCTTACCCCAAGCTTTTAAAAATA
 15 TTTTACCTGCTCTAGGGAATGAATTTATTACAATTATTAAAGATAGTGCTCT
 CCTTCAAACCTATTGGTGTCTATGGAATTATGGAACGGAGCACAATCAGTTGT
 AACGGCTACTTACTCACCAGTTGCACCGTTATTATTTGCAGCATTTTACTAT
 TTAATGTTGACAACGATTCTCTCAGCTTTGTTAAAACAAATGGAGAAATAT
 CTTGGGAAAGGGGTAAAAATAGATGGTTGA
 20 MSMNFSFLPQYWSYFNYGVMVTIMISTCVVFFGTIIGVLIALVKRTNLHFLTIL
 ANFYVWVFRGTPMVVQIMIAFAWMHFNNLPTISFGVLDLDFTRLLPGIIISLNS
 GAYISEIVRAGIEAVPSGQIEAAYS LGIRPKNTLRYVILPQAFKNILPALGNEFITI
 IKDSALLQTIGVMELWNGAQSVVTATYSPVAPLLFAAFY YLMLTTILSALLKQ
 MEKYLKGKGVKIDG*

25 Sequence description:

- 30 A] Length: 699 bp - 233 aa (full length gene)
 B] Shine-Dalgarno sequence preceded the 'ATG'
 start codon. Possesses a potential leader peptide
 sequence.

ID-113

35 Clone 2-7

40 ATGAAAGACCTATTACGAAATAGTCTAGAGCAAAGTGGAAATTTAAGTTT
 TCAAGATATGATTTTACATATTCTTGTAGCAGCTTTATTGAGTGTTAGTTATT
 TATGTTTCCTATGCTTATACGCATAGTGGAAGTGCCTATAGTAAAAAGTTT
 AATGTTTCATTAATGACATTGACGGTCTTGACTGCAACAGTAATGACCGTT
 ATTGGTAATAATGTAGCCTTGTCATTGGGTATGGTCGGTGCCTTGTCAGTT
 GTTCGTTTTAGGACAGCCATAAAAGATTCAAGAGATACAGTTTATATTTTT
 TGGACCATAGTTGTTGGTATCTGTTGTGGTGTGGTGACTATGTGGTAGCT

GCATTAGGAAGTAGCGTTATCTTTATCTTATTATGGGTTATGGGACGTGTT
 AAAAACGAGAATCGTATGTTATTGATTGTGAAGTGCATAGAACACTAGA
 AGTTGATTTAGAAAGGAATTTTCTTCCAATATTTTGACGGAAAAGCTGTTCA
 GCGTGTTAAAAATTCAACAATACTATTGAAATGATTTTCGAAATCTC
 5 TAGAAAAGATTACGATAAGCAACTCCATGTAGATAATCAGTTAACTGAAA
 AAGTGTACCAATTGGGAAATATTGATTATTTCAACATTGTTAGCCAAAGCG
 ACGAAATCAATGGGTAG

MKDLLRNSLEQSGNLSFQDMILHILVAALLSVVIYVSYAYTHSGTAYSKKFNV
 10 SLMTLTVLTATVMTVIGNNVALSLGMVGALS VVRFRRTAIKDSRDTVYIFWTIV
 VGICCGVGDYVVAALGSSVIFILLWVMGRVKNENRMLLIVKCDRTLEVDLEGI
 FFQYFDGKAVQRVKNSTTNTIEMIFEISRKDYDKQLHVDNQLTEKVYQLGNID
 YFNIVSQSDEING*

15 Sequence description:

A] Length: 678 bp - 226 aa (full-length gene)
 B] ATG start codon is preceded by a Shine-
 20 Dalgarno sequence-Possesses a potential leader
 peptide sequence

25 ID-114

Clone 2-8

30 AAAAATTCATTTTAGATTTCATTTTACGACTATATACTCAGAAGTACCAAAC
 CTAATCCAAGGTTTGAAAAAAGAAAGAAGGAAGTCAGTATGACAACTAT
 AAAACAACCTTAAAGATGAGGCTATACGTGTTGAAGAGACAACAAAAGA
 ATCATTTTACGATGTTGATATTGCCTTGTTTTCAGCTGGTGGATCTATTTCA
 GCAAAGTTCGCTCCTTATGCAGTAAAGTCTGGAGCAGTTGTAGTAGATAAC
 35 ACGTCATATTTTCGTCAGAATCCTGATGTTCCACTAGTTGTTCCCTGAAGTAA
 ATGCTCATGCCATGATTGGTCATAATGGTATCATAGCTTGTCCCAATTGTTT
 TACTATTCAAATGATGATTGCTTTAGAGCCCATTCGTCAAAAATGGGGGAT
 AGAGCGTGTTATAGTTTCCACCTATCAAGCTGTTTCGGGTTACAGGTGCACG
 TGCTGTTGAAGAACTAAGGAACAGTTGAGACAAGTTTT

40 KFILDSFYDYILRSTKPNPRFEKRKKEVSMTNYKNNFKDEAIRVEETTKESFYD
 VDIALFSAGGSISAKFAPYAVKSGAVVVDNTSYFRQNPDPVPLVVPEVNAHAMI
 GHNGIIACPNCSTIQMMIALEPIRQKWGIERVIVSTYQAVSGSGARAVEETKEQ
 LRQV

Sequence description:

- 5 A] Length: 499 bp - 165 aa (partial sequence)
 B] N-terminus has yet to be determined

ID-115

10

Clone 2-9

15 ATGACAAATGAATTGATAATGCAAGCTTTTGAGTGGTATTTACCTAGTGAT
 GGGAATCACTGGAAGAAATTAGAGGAGTCTATATCAGACCTTAAAAAACT
 TGGAATTAGTAAAATCTGGTTACCACCAGCATTTAAGGGAACTAGCAGTG
 ATGATGTAGGATATGGTGTTTATGATCTCTTTGATTTAGGAGAATTTGACC
 AGAATGGAACAATTAGAACAAAATATGGTAGGAAAGAAGAGTATCTAAA
20 GCTTATTAAGTCGTTAAAGGCAAATGGCATTAAACCGTTTGCAGATATCGT
 TCTTAACCATAAAGCCAATGGTGATCATAAAGAAAAATTTCAAGTCATCA
 AAGTCAATCCTGAAAATCGTCAAGAAGCATTAAGTGAACCCTATGAGATT
 GAAGGATGGACGGGATTTGATTTCCCAGGTAGACAGGGTGAGTACAATGA
 TTTT
25 MTNELIMQAFEWYLPDGNHWKKLEESISDLKKLGISKIWLPPAFKGTSSDDV
 GYGVYDLFDLGEFDQNGTIRTKYGRKEEYLKLIKSLKANGIKPFADIVLNHKA
 NGDHKEKFQVIKVPENRQEALSEPYEIEGWTGFDFPGRQGEYNDF

30

Sequence description:

- 35 A] Length: 456 bp - 152 aa (partial sequence)
 B] ATG start codon is preceded by a Shine-
 Dalgarno sequence, no leader peptide sequence.

ID-116

40

Clone 2-10

ATGGAGGTTCTTATGAAGAAAGTGTTAGTAAGTAGTCTTTTGGTTTTAGGG
ATTACGATAACGTTACAACCAGTAGTTGAGGCTAAGGGGCCAAAAGTAGC
TTATACACAAGAGGGAATGACTGCTCTTTTCGGACACAAATAAAGATAAAG
5 TCACTACTATTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGC
CGATTACTGTTAGTTTTGATATTGATGATACACTGCTTTTTCAGTAGTCAATA
TTTTCAATATGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCAT
AAACAAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCAT
TCCCAAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAG
10 ATAAAATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAG
GGCGAGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTTGTA
CCATCTGAT

MEVLMKKVLVSSLLVLGITITLQPVVEAKGPKVAYTQEGMTALSDTNKDKVT
15 TISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQYGKEYVTPGSFDFLHKQKFW
DLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGSMYKEGEVDKTA
KALAKDFKFVPSD

20 Sequence description:

A] Length: 516 bp - 172 aa (partial sequence)
B] ATG start codon is preceded by a Shine-
Dalgarno sequence, Possesses a leader peptide
sequence.

25

ID-117

30 Clone 2-17

ATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGACAAAGATTAGTCGT
35 CGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTATTATTGTCAGGATGT
GTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAACTCCGAATGTTTCAGC
GCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTTAATTCTTTTCGTAGTTT
AGATCAGCTAACTAGCTTTAAAGAGATTCTTTGGGTTATTGGTCAAAATGT
AGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGTTACTATCCCTAAAG
CCAAGTTTACGGAAATATAAAAGCGTTATATTACTTGCTTTCTTGATGTCTC
40 TTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTTAATAGATGCTAATCG
GGTTTTTGAAATCGACGATCTATGGACAAATACCTTAGGCGGTCTTTTCGC
CCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCTAACTATTAGAAAATG
A

MLKRLFTEDGELTKISRRFVWMLVVIYCLIIVRMCFGPQIMIEGVSTPNVQRFG
 RIVALLVPFNSFRSLDQLTSFKEILWVIGQNVVNILLFPLIIGLLSLKPSLRKYK
 SVILLAFLMSLFIECTQVVLDILIDANRVFEIDDLWTNTLGGPFALWSYRNIKG
 WLLTIRK*

5

Sequence description:

- 10 A] Length: 516 bp - 172 aa (full-length gene)
 B] ATG start codon is preceded by an Shine-
 Dalgarno sequence. Possesses a potential leader
 peptide sequence. C-terminus need further
 confirmation.

15

ID-118

20 Clone 3-3

25 ATGAAAAAGCTTACTTTTATTTGGGATTTAGATGGGACATTAATAGATTTCG
 TATGTACCAATTATGGAAGCTCTTGAAGAAACCTATCGTCATTTTGGCTTA
 ATATTTGATAAAGAATTAATCCATGAATATATTTTACAGGAATCAGTGGGG
 CAATTATTGGTAAACCTTTCAGAGGAAGAGCAAATACCTCATGAAAAACT
 GAAAGCATATTTTACAAAAGAACAAGAAAGTCGAGATTCTAAAATACATT
 TAATGCCATATGCAAAAGAGATTTTAGAATGGACCAAAGAACAAGATATT
 CCCAATTTTATGTATACACATAAAGGAGCAAGTACGCATTTCAGTGTGGA
 ACCTTGCAGATCTCTCATTATTTTGATGAAATTTTAACTGGTGTTCGGGAT
 30 TCGAGCGAAAACCACATCCACAAGGGATTAATTATTTAGTTAAACGATATT
 CTTTAGATAAATCAATGACTTATTACATAGGAGATCGTCCACTAGATTTGG
 AGGTTGCTCAAAATGCTGGTATAAAATCCATAAACTTAAGGTTAGAGAATT
 CCAAAGAAAACCTATAATATTTCAAGTCTCAAAGATATAATATCACTTGATT
 TCACTCGTTTGGATTAA

35

 MKKLTFIWDLDGTLIDSYPIMEALEETYRHFGLIFDKELIHEYILQESVGQLL
 VNLSEEEQIPHEKLKAYFTKEQESRDSKIHLMPYAKEILEWTKEQDIPNFMYTH
 KGASTHSVLETLQISHYFDEILTGVSGFERKPHPQGGINYLVKRYSLDKSMTYYI
 40 GDRPLDLEVAQNAGIKSINLRLNSKENYNISSLKDIISLDFTRLD*

Sequence description:

- A] Length: 627 bp - 209 aa (Possible Full-length gene)

B] ATG start codon is preceded by an possible
Shine-Dalgarno sequence. No obvious leader
peptide sequence.

5

ID-119

10

Clone 3-7

15

20

25

30

35

40

ATGGAAAAAGAAAAAAATTAGGTCTTTTACCACTAACAATGCTTGTTCATT
GGCTCTCTTATCGGTGGCGGAATCTTTGATTTAATGCAAAATATGAGTTCC
AGAGCCGGTTTGGTACCAATGCTTATTGCTTGGGTAATTACTGCTATCGGG
ATGGGAACTTTCGTTTTAAGTTTTCAAAATTTATCTGAAAAAGGCCGGAC
CTAACAGCTGGAATCTTTAGTTACGCTAAAGAGGGGTTTGGAACTTTATG
GGATTTAACTCTGCATGGGGTTATTGGTTATCAGCTTGGCTTGGAAATGTT
GCCTACGCTGCACTCTTATTCAGTTCAGTTCGGTTATTTCTTTAAATTCTTTG
GTAATGGAAATAATATCATCTCAATTATTGGAGCAAGTATAGTTATTTGGG
TTGTCCATTTCTTAATTTTAAGAGGTGTTAATACAGCTGCATTTATTAATAC
CGTAGTTACCTTTGCAAAATTAGTACCTGTTATTATTTTCTTAATTTACAGCG
TTATTAGCTTTCAAATTTAACATTTTTAGTCTTGATATCTGGGGAAATGGAT
TACATCAATCAATTTTCAACCAAGTCAATTCAACTATGAAAACCGCTGTTT
GGGTATTTATTGGTATTGAGGGCGCCGTTGTCTTCTCAGGTCGTGCTAAAA
AACACTCTGATATTGGTAAAGCAAGTATCCTAGCATTATTCAGTATGATTT
CACTTTATGTATTGATTTCTGTTTTATCACTTGGTATCATGTACAGTCCAGA
ACTTGCAAACTTAAAAACACCAGCTATGGCTTACGTTCTAGAAAAAGCTGT
TGGTCACTGGGGTGCTATCTTAGTTAACCTTGGTGTATCATTTTCAGTATTT
GGCGCTATTCTTGCTTGGACTTTATTTGCAGCAGAATTACCATATCAAGCT
GCTAAAGAAGGTGCTTTTCCTAAATTTTTTGCAAAAGAAAAATAAAAAACAA
AGCTCCAATCAACTCACTCTTAGTCACTAATCTTTGTGTACAAGCATTCTTA
ATCACGTTCTTATTCACACAAAGTGCTTATCGTTTTGGTTTTCGCATTAGCAT
CATCTGCTATCTTAATTCCTTATGCTTTTACAGCACTATATCAATTACAATT
CACACTCCGTGAGGATAAGTCAACTCCAGGACATCAAAAGAATTTAATTA
TCGGTATCCTCGCTACAATCTATGCTGTTTACCTTATCTACGCTGGTGGTTT
TGATTACTTACTTTTGACAATGATTGCTTATACTCTAGGTATGATTCTCTAT
ATTAAAATGAGAAAAGATGACAAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCAGTGTGAAATTGTTATCC

MEKEKKLGLLPLTMLVIGSLIGGGIFDLMQNMSSRAGLVPMLIAWVITAIGMG
TFVLSFQNLSEKRPDLTAGIFSYAKEGFGNFMGFNSAWGYWLSAWLGNVAY
AALLFSSLGYFFKFFGNGNNIISIIGASIVIWVVHFLILRGVNNTAAFINTVVTFAK
LVPVIIFLISALLAFKFNIFSLDIWGNGLHQSFNQVNSTMKTAVVWVFIGIEGAV

VFSGRAKKHSDIGKASILALFTMISLYVLISVLSLGIMSRPELANLKTPAMAYV
LEKAVGHWGAILVNLGVIIISVFGAILAWTLFAAELPYQAAKEGAFPKFFAKEN
KNKAPINSLLVTNLCVQAFLITFLFTQSA YRFGFALASSAILIPYAFTALYQLQF
TLREDKSTPGHQKNLIIGILATIYAVYLIYAGGFDYLLLTMIAYTLGMILYIKMR
5 KDDKLGVMVIAVSSVKLLS

Sequence description:

10

A] Length: 1356 bp - 452 aa (partial sequence)
B] ATG start codon is preceded by an possible
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

15

ID-120

20

Clone 3-8

ATGAAATTTGAAAAACGGCAGGTCTATTATGTTGTCATAACATTTGCTATT
25 TGCTATGCTATACAGGCTTATTGGGGAGCTGTTTCTAATATTTTAACTACGC
TTCATAAGGCAATATTTCTTTTTTGATGGGAGCTGGAATTGCCTATATTAT
TAATATTGTAATGTCAGTCTATGAGCGATTATATATAAAGCTTTTTAAAGG
ATCTAGACTATTAATGGCAATCAAGCGTAGTGTTTCTATGATTTTATCCTAT
GCAACTTTTATTGGTTTAATTGTCTGGCTATTTTCAATTGTCATTCCAGATT
30 TGATTTCTAGTTTGAGTTCTTTATTGGTTATTGATACCGGAGCACTTGCTAA
ATTGGTTAATAATCTCAATGAAAATAAACAATTTCTGAGGCTTTAAATTA
TATGGGAACAGATAAAGACTTAGTTTCTACTTTAAGTGGTTATAGCCAGCA
GATTTTGAAGCAAGTTTTATCTGTTTTTAACAAATTTACTAACCTCAGTTTCC
TCTATTGCGGCAACACTTCTGAATGTTTTTGTAGTTTTATTTTTTCAATTTA
35 CGTTTTGGCAAACAAGGAGCAGTTGGGACGTCAATTTAATTTGTTAATTGA
TACCTATTTAGGTTCAACAGGCAAAACATTCCATTACGTTTCGTCATATCCTT
CATCAACGTTTCCATGGTTTTTTTTGTAAGCCAAACTTTAGAAGCTATGATTT
TAGGAAGTTTGACGGTTATTGGTATGTTGATCTTCCAATTTCTTATGCTTT
AACAGTTGGGGTTTTAGTTGCTTTTACAGCTCTAATACCGGTTGTGGGAGC
40 CTACATTGGTGTTACAATCGGTTTCATCTTAATTGCTACTGAATCGCTTACT
GAAGCATTCTTGTTTGTCTTTTCTTGATCCTTTTACAACAATTTGAGGGAA
ATGTCATTTATCCGAAAGTTGTCGGTGGATCGATTGGACTGCCTTCTATGT
GGGTTTTAATGGCTATTACTATCGGAGGTGCTTTATGGGGGATCTTAGGCA

TGTTACTTGCTGTTCTGTTGCAGCTACTATCTATCAGATTGTAAAAGATCA
TATTATCAAGCGACAAACGCTTAGAAATCGTGCACGAACCTATCGTTAA

5 MKFEKRQVYYVVITFAICYAIQAYWGAVSNILTTLHKAIFPFLMGAGIAYIINI
VMSVYERLYIKLKFSGSRLMAIKRSVSMILSYATFIGLIVWLFSIVPDLISLSS
LLVIDTGALAKLVNNLNENKQISEALNYMGTDKDLVSTLSGYSQQILKQVLSV
LTNLLTSVSSIAATLLNVFVSFIFSIYVLANKQLGRQFNLLIDTYLGSTGKTFH
YVRHILHQRFHGFFVSQTLEAMILGSLTVIGMLIFQFPYALTVGVLVAFTALIP
10 VVGAYIGVTIGFILATESLTEAFLFVLFLILLQQFEGNVIYPKVVGGSIGLPSM
WVLMAITIGGALWGILGMLLAVPVAATYQIVKDHIKRQTLRNRARTYR*

Sequence description:

15 A] Length: 1134 bp - 378 aa (full-length gene)
B] ATG start codon is preceded by an typical
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

20

ID-121

Identical to ID-68, as described in WO 00/06736

25

ID-122

30

Clone 3-16

35 GTGATTACAATTAAAAAGGAATCTGTTATCAAACCTATTGAAGTATGCTTTT
GGCATTATAATGGGATTTATTATCTTAGCTATTGTAATAGGTGGGCTCCTA
TTTGCATACTACGTTAGTCGTTCTCCGAAATTAACCGATCAAGCTTTAAAA
TCCGTTAACTCTAGTTTGGTTTATGATGGTAATAATAAACTTATTGCCGATT
TAGGCTCAGAAAAGCGTGAAAGTGTTAGTGCGGATAGCATTCCACTAAAT
TTGGTTAACGCTATCACTTCTATAGAAGATAAACGTTTCTTTAAACATAGA
40 GGTGTCGATATTTATCGTATTTTAGGTGCAGCTTGGCATAACCTTGTTAGTA
GTAATACGCAAGGTGGTTCAACCCTTGATCAACAGTTGATTAACTGGCTT
ACTTTTCTACCAATAAATCTGACCAAACGTTAAACGTAATCACAGGAA
GTTTGGCTTGCGCTTCAAATGGAGCGTAAATACACCAAAGAAGAAATTCTT
ACTTTCTATATTAATAAAGTTTATATGGGAAATGGGAATTATGGTATGAGA

ACAACAGCTAAATCATACTTTGGTAAAGACCTAAAGGAATTATCTATTGCA
CAACTTGCTTTGCTCGCTGGTATTCTCAAGCACCTACACAATATGACCCTT
ATAAAAACCCAGAATCTGCTCAAACAAGACGTAATACCGTTCTTCAGCAG
ATGTATCAAGATAAAAACATTTCTAAAAAGGAATACGACCAAGCTGTTGC
5 AACTCCAGTAACTGATGGCTTAAAAGAATTAAAGCAAAAATCTACTTATCC
AAAATATATGGATAACTACTTAAAACAAGTTATTAGTGAAGTTAAACAAA
AACTGGTAAAGATATCTTTACTGCTGGGCTAAAAGTGTATACTAATATCA
ACACTGATGCACAAAAACAATATATGACATCTACAACAGTGATACTTAC
ATCGCTTATCCAAACAATGAATTACAAATAGCATCTACCATCATGGATGCG
10 ACTAATGGTAAAGTCATTGCACAATTAGGCGGGCGTCATCAGAATGAAAA
TATTTTCAATTTGGGACAAATCAATCTGTCTTAACAGACCGCGATTGGGGTTC
TACAATGAAACCTATCTCAGCTTATGCACCTGCTATTGATAGTGGTGTCTA
TAATTCAACAGGTCAATCATTAACGACTCAGTTTACTACTGGCCTGGTAC
TTCTACTCAACTATATGACTGGGATCGTCAATATATGGGTTGGATGAGTAT
15 GCAGACCGCTATTCAACAATCACGTAACGTCCCTGCTGTCAGAGCACTTGA
AGCCGCTGGATTAGACGAAGCAAAATCTTTCCTTGAAAAATTAGGCATAT
ACTATCCAGAAATG

MITIKKESVIKLLKYAFGIIMGFILAIIVIGLLFAYYVSRSPKLTQALKSVNSS
20 LVYDGNKLIADLGSEKRESVSADSIPLNLVNAITSIEDKRFFKHRGVDIYRILG
AAWHNLVSSNTQGGSTLDQQLIKLAYFSTNKSQTLKRKSQEVWLALQMER
KYTKEEILTFYINKVYMGNGNYGMRTTAKSYFGKDLKELSIAQLALLAGIPQA
PTQYDPYKNPESAQTRRNTVLQQMYQDKNISKKEYDQAVATPVTDLKELK
QKSTYPKYMDNYLKQVISEVKQKTGKDIFTAGLKVYTNINTDAQKQLYDIYN
25 SDTYIAYPNNELQIASTIMDATNGKVIAQLGGRHQNENISFGTNQSVLTDRDW
GSTMKPISAYAPAIDSGVYNSTGQSLNDSVYYWPGTSTQLYDWDQRQYMGWM
SMQTAIQSRNVPVRALEAAGLDEAKSFLEKLGIIYPEM

30 Sequence description:

A] Length: 1386 bp - 462 aa (partial sequence)
B] GTG start codon is preceded by an
35 typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

40 ID-123

Clone 3-17

ATGGCTAATGTATATGATTTAGCAAATGAATTAGAACGTGCTGTTCGTGCT
 TTACCAGAATACCAAGCAGTTTTAACTGCAAAAGCAGCTATTGAAAATGA
 TCGGATGCACAAGTGCTTTGGCAAGACTTTTTGGCTACCCAATCAAAAGT
 TCAAGAAATGATGCAATCTGGCCAAATGCCAAGTCAAGAAGAACAAGATG
 5 AAATGTCTAAACTTGGGGAAAAAATTGAATCCAATGACCTTTTAAAAGTTT
 ATTTTGACCAACAACAACGGTTGTCTGTCTATATGTCTGATATCGAAAAAA
 TTGTCTTTGCACCCATGCAGGACTTGATGTAA

MANVYDLANELERAVRALPEYQAVLTAKAAIENDADAQVLWQDFLATQSK
 10 VQEMMQSGQMPSQEEQDEMSKLGEKIESNDLLKVYFDQQRLSVYMSDIEKI
 VFAPMQDLM*

Sequence description:

15 A] Length: 336 bp - 112 aa (full length sequence)
 B] ATG start codon is preceded by an
 typical Shine-Dalgarno sequence. No obvious
 potential leader peptide sequence.

20

ID-124

25

Clone 3-26

ATGGCAGAAATCACAGCTAAACTTGTAAGAATTGCGTGAAAAATCAGG
 30 TGCAGGCGTTATGGACGCTAAAAAAGCATTAGTAGAACTGATGGTGACC
 TTGATAAAGCGATTGAATTACTTCGCGAAAAAGGTATGGCTAAAGCAGCT
 AAAAAAGCAGACCGTGTTGCTGCTGAAGGTTTAACAGGTGTTTATGTTGAT
 GGTAACGTTGCAGCAGTTATTGAAGTTAA

35 MAEITAKLVKELREKSGAGVMDAKKALVETDGDLDKAIELLREKGMAKAAK
 KADRVAEGLTGVYVDGNVAIVIEV

Sequence description:

40 A] Length: 230 bp - 76 aa (partial sequence)
 B] ATG start codon is preceded by an
 typical Shine-Dalgarno sequence. No obvious
 potential leader peptide sequence.

ID-125

5

Clone 3-33

10 ATGATAAAAAACCTGTTATTAACAGGTTTTTTATCATTTAATGACGGAAAA
CTGGACACAAATTATTTTTCTTGTATAATTAAATATATTATTTCTTATCAGG
AGGTTATGATGACATTAGAGAAACGATTTAA

15 MIKNLLLTGFLSFNDGKLDTNFYFSCHIKYIISYQEVMMTLEKRF

Sequence description:

20 A] Length: 134 bp - 44 aa (partial sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. Possible
potential leader peptide sequence.

25

ID-126

Clone 3-41

30

ATGAAAAATAATAAAAAATAATGGTTTTCTGAAAAATTCCTTTATTTACATA
TTATTGATTATTGCGGTTATTACAACCTTTCAATACTATTTAA

35 MKNNKNNGFLKNSFIYILLIHAVITTFQYYL

Sequence description:

40 A] Length: 94 bp - 31 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

ID-127

5

Clone 3-42

10 ATGTTAGATATTATCTTATCCGGAATTTTCGCAAGGATTACTTTGGTCAATTA
TGGCAATTGGCGTGTTTATCACTTTTCGTATCTTAGACATAGCCGATCTCTC
TGCAGAAGGGGCTTCCCTATGGGGGCTGCAGTTTGCGCCTTATGTATCGT
TAA

15 MLDIILSGISQGLLWSIMAIGVFITFRILDIADLSAEGAFPMGAAVCALCIV

Sequence description:

20 A] Length: 158 bp - 52 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

25

ID-128

30 Clone 3-43

35 ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAATAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTTAA

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTAVILSLF

40 Sequence description:

A] Length: 161 bp - 53 aa (full-length gene)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential

leader peptide sequence.

ID-129

5

Clone 3-44

10 GTGGTAAGTAAATTGAGTTTAAACAACGATTTTTGCATTGCTATTTTCATCA
ATGCTAATTTACGCAACACCTCTTATCTTTACAAGTATTGGGGGAACCTTC
TCTGAACGTGGTGGTATCGTCAACGTTGGTTTAGAAGGAATTATGGTAATT
GGAGCTTTCTCAGGCGTTGTATTTAA

15 MVSKLSLTTIFALLFSSMLIYATPLIFTSIGGTFSEKGGIVNVGLEGIMVIGAFSG
VVF

20 Sequence description:

A] Length: 179 bp - 59 aa (partial sequence)

B] GTG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

25

ID-130

30

Clone 3-46/47

35 ATGAGAATTATTGCAATAACTGAAAAGGTTATAAAAGAACTGTTTCGTGAT
AAAAGAACACTTGCTATGATGTTTTTAGCACCTATTTTAATTATGTTTTTGA
TGAATGTTATGTTTTCTGCGAATAGTAATACAAAAGTTAAGATTGGAACATA
TTAACGTTAACACGAAGGTCGTTTCAAATTTAGATAATATTAAGCATATTC
AAGTGAGATCATTTAAATTTAACTCATCTGCTAAAAAAGCACTCAAATCAA
ATAAAATTGATGCTCTTATTTTCGGAGGACAATAAATCTTATACTGTCTTCT
40 ATGCGAATACAGATTCTTCAAAGACGACTTTAACAAGACAAGCTTTTAAA
ACCGCTGTTAATAACAATGAACAGTAAGGAACTGATTTTCGCAAGTTAAAATT
TTAGCTAATAAGAATCCGAACTAGCACAATCCTTACAACTCGCTCCAAA
TATATCAAAGAAAAATATAATTACGGAAATAAAAAATACAGGCTTTTTTGC
AAAAATGATACCAATACTAATGGGATTTATGGTCTTCTTCTTGGTTTTT

5 MRIIAITEKVIKELFRDKRTLAMMFLAPILIMFLMNVMFSA NSNTKVKIGTINV
 NTKVVSNLDNIKHIQVRSFKFNSSAKKALKSNKIDALISEDNKS YTVFYANTDS
 SKTTLTRQAFKTA VNTMNSKELISQVKILANKNP KLAQSLQTRSKYIKEKYN Y
 GNKNTGFFAKMIPILMGFMVFFLVF

Sequence description:

10 A] Length: 558 bp - 186 aa (partial sequence)
 B] ATG start codon is preceded by a
 possible Shine-Dalgarno sequence. Potential
 leader peptide sequence. C-terminus has yet to be
 determined.

15

ID-131

20

Clone 3-48

25 GTGATTATCGTTATGAGTAAACATCAAGAAATTTTGGAGTACCTAGAAAAT
 TTAGCTGTTGGTAAGAGGGTTAGTGTACGCAGTATTTCAAATCATTTAA
 MIIVMSKHQEILEYLENLAVGKRVS VRSISNHL

30 Sequence description:

35 A] Length: 100 bp - 33 aa (partial sequence)
 B] GTG start codon is not preceded by a
 obvious Shine-Dalgarno sequence. No obvious
 leader peptide sequence.

40 ID-132

Clone 2-c53

ATGTATAGAGAAATTACCGCTGTCGAACACGATCGCTTTGTGAGCGAATCC
 AACCAAACAAACCTACTTCAATCTCTTAATTGGCCCAAAGTAAAAGACAA
 CTGGGGTAGTCAATTACTTGGCTTTTTTGGACGGTGAAACCCAAATTGCCAG
 CGCTAGTATTCTCATCAAATCACTTCCTCTTGGCTTCTCCATGCTGTATATT
 5 CCGCGTGGACCAATCATGGATTACTCCAATCTAGATATTGTAACCTAAGGTC
 CTTAAGGACCTTAAAGCTTTTGGCAAAAAACAAAGAGCTCTCTTTATCAAG
 TGTGATCCTCTCATCTATT

10 MYREITAVEHDRFVSESNQTNLLQSLNWPVKVDNWGSQLLGFFDGETQIASA
 SILIKSLPLGFSMLYIPRGPIMDYSNLDIVTKVLKDLKAFGKKQRALFIKCDPLI
 Y

15 Sequence description:

20 A] Length: 326 bp - 108 aa (partial sequence)
 B] ATG start codon is preceded by an obvious
 Shine-Dalgarno sequence. No obvious leader
 peptide sequence.

25 ID-133

Clone 2-c59

30 ATGGACAAGAAAAAATCTTAGTAACGGGTATTGTGCCTAAAGAAGGTCT
 AAGAAAGCTTATGGACCGATTTGATGTTACTTATTCAGAAGATCGCCCATT
 TTCACGTGACTATGTGTTAGAGCATTTATCTGAATATGACGGATGGTTACT
 CATGGGACAAAAAGGTGATAAAGAGATGATTGATGCAGGTGAAAACCTAC
 35 AAATTATTTCTTT

MDKKKILVTGIVPKEGLRKLMDRFDVTYSEDRPFSRDYVLEHLSEYDGWLLM
 GQKGDKEMIDAGENLQIIS

40 Sequence description:

A] Length: 215 bp - 71 aa (partial sequence)

B] ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

5

ID-134

10

Clone 2-c62

15

ATTTTCGAAAGATGACTACCAAAATATTAGTTTTGGACAGGATCCAGAAGTT
GTTGATTATGCTGGTCTGTTTGAAAAACGCCGTCCAGTTTTAGAAAAAGCA
GTTAAAAATTTCTTGCAAGAAGAGAGAGCTACGAGAATGCTATCTGATTTT
TTGCAAGAAGAAAAATGGGTAACCTGATTTTGCTGAATTTATGGCGATCAA
AGAACATTTTGGTAATAAGGCGCTTCAAGAATGGGATGACAAGGCTATTA
TACGCCGCGAAGAAGAAGCCTTAGCAGGATATCGTCAAAAGCTTAGTGAA
GTGATAAAATATCATGAAGTAACGCAATATTTCTTTTACAAACAATGGTTT
GAGTTAAAAGAATATGCTAATGATAAAGGGATTCAAATTATCGGTGATAT
GCCAATCTACGTTTCTGCCGATAGTGTAGAAGTTTGGACAATGCCTGAACT
GTTT

20

25

ISKDDYQNI SFGQDPEVVDYAGLFEKRRPVLEKAVKNFLQEERATRMLSDFLQ
EEKWVTDFAEFMAIKEHFNGKALQEWDDKAIIRREEEALAGYRQKLSEVIKY
HEVTQYFFYKQWFELKEYANDKGIQIGDMPYVSADSVEVWTMPFLF

30

A] Length: 459 bp - 153 aa (partial sequence)
B] More sequencing is required to determine the
N- and C-termini
enzyme). - *Streptococcus pneumoniae* (63%)

35

ID-135

Identical to ID-108 described in WO 00/06736

40

Clone 2-c63

ID-136

Clone 2-c66

5

30

35

40

ID-137

Clone 2-c67

5

TTGAGAGTTTATGAAAATAAAGAAGAGTTGAAAAAAGAAATAAGTAAAAC
ATTTGAGAAATACATTATGGAATTTAATAA
TATTCCAGAGAATCTAAAAGATAAAAGAATTGATGAAGTTGATAGAACTC
10 CAGCAGAAAACCTTTCTTATCAGGTTGGCT
GGACCAACTTGGTTCTTAAATGGGAAGAAGATGAAAGAAAGGGACTTCAA
GTAAAAACACCATCGGATAAATT
MRVYENKEELKKEISKTFEKYIMEFNNIPENLKDKRIDEVDRTPAENLSYQVG
15 WTNLVLKWEEDERKGLQVKTPSDKF

Sequence description

20

A] Length: 234 bp - 78 aa (partial sequence)
B] TTG start codon is preceded by a
potential Shine-Dalgarno sequence. No obvious
leader peptide sequence.

25

ID-138

30

Clone 2-c70

ATGTCAAAGTTTGATAGTCAGAAAATAATTACTCCGATTATGAAGTTTGTC
AATATGCGAGGGATTATTGCACTCAAAGATGGCATGCTAGCAATTTTACCA
35 CTAACAGTTGTTGGGAGTCTCTTTTAAATATTAGGGCAGCTTCCATT

MSKFDSQKIITPIMK FVNMRGIIALKDGMLAILPLTVVGSFLILGQLPF

40

Sequence description

A] Length: 150 bp - 50 aa (partial sequence)
B] ATG start codon is preceded by a potential
Shine-Dalgarno sequence. Possesses a potential

leader peptide sequence.

ID-139

5

Clone 2-c71

10 GAGACCACTTCATCAGTTAAACCAGCAGGAATTGACCGTATCAATCATACC
TCAACACCCCCGAAGAAAAC TACCCCAACATTGCAACGACGCATAGCTT
CAAAGATCGTTGTGATACTTTAGAAAGAATTCACAATGAAGACATTGATGT
TTGTTCTGGATTCATTTGTGGTATGGGAGAGAGCGATGAGGGGCTCATCAC
ATTAGCTTTCAGACTAAAAGA ACTGAACCCCTATTCTATCCCTGTCAATTTT
15 TTACTTGCTGTTGAAGGAACACCTCTTGGAATAATAACTATTTGACTCCC
ATTAAATGCTTAAAAATTATGGCCATGTTGCGTTTTGTTTTCTTTCAAGG
AATTAAGATTAAGTGCTGGACGGGAGGTCCATTTTGAGAATTTTGAATCAT
TAGTCACCTTACTTGTTGACTCAACTTTTTTGGGAAATTACCTAACAGAGG
GGGGTCGCAATCAACATACCGATATTGAATTCTTGGAATAATTACAATA
20 AATCATACTAAAAAGGAATTAATTT

ETTSSVKPAGIDRINHTSTPPKKTPNIATTHSFKDRCDTLERIHNEIDVCSGFI
CGMGESDEGLITLAFRLKELNPYSIPVNFLLA VEGTPLGKYNYLTPIKCLKIMA
MLRFVFPFKELRLSAGREVFHFENFESLVTL LVDSTFLGNYLTEGGRNQHTDIEF
25 LEKLQLNHTKKELI

Sequence description:

30

A] Length: 535 bp - 178 aa (partial sequence)
B] N- and C-termini require verification

35

ID-140

Clone 2-c73

40

ATGCCGGTTTGGACTGCACAGTCTATTCCAAAGGCATTTT TAGAAAAGCAT
AATACTAAGGAAGGCACCTGGGCAAACTAACCATTCTAAGTGGTTCTTTA
GTATTTTACCAGTTATCTCTGATGGAGAGGAAATCTCGCGGCATATTTT

5 GATGCTAGTAGTGATATTCCTTTTGTGATCCACAAGTCTGGCATAAAGTT
TCGCCGAATAGTCCAGACTTAAGTTGCTATCTAACTTTTTACTGCCAAAAA
GAAGATTACTTCCATAAAAAATATGGTCTCACGCGCACACATTCTGAGGTT
ATCGCCAGTGCACCTCTCTTATCTGAGAAGAGTAATATATTAGACCTTGGG
TGTGGTCAAGGGCGAAACTCACTTTATTTATCGCTGCTGGGACATCAAGTG
ACTTCTGTGCGATTCAAACGGACAGAGCCTTGTAGCTTTAGAAAATATGGCA
TTAGAAGAAGAGCTTCCTTACAATATAAAAAGGTATGATATTAATACTACT
GCTATTGAAGGGCACTATGATTTTATTTTATCAACTGTGGTATTTATGTTTT
T

10 MPVWTAQSIPKAFLEKHNTKEGTWAKLTILSGSLVIFYQLSPDGEEISRHIFDAS
SDIPFVDPQVWHKVSPNSPDLSCYLTIFYCQKEDYFHKKYGLTRTHSEVIASAP
LLSEKSNILDLGCGQGRNSLYLSLLGHQVTSVDSNGQSLVALENMALEEELPY
15 NIKRYDINTTAIEGHYDFILSTVVFMF

Sequence description:

20 A] Length: 563 bp - 187 aa (partial sequence)
B] N- and C-termini require verification

25 ID-141

Clone 2c76

30 ATGACAAAGCAAATAATTGCCATTTGGGCTGAAGATGAAGACCATTTGAT
TGGAGTTAATGGCGGTTTACCATGGAGGCTTCCTAAAGAGTTACATCACTT
CAAAGAAACGACCATGGGGCAGGCTTTGCTTATGGGACGAAAGACCTTTG
ATGGAATGAACCGTCGTGTTTTACCTGGTAGAGAGACAATCATCTTAACAA
35 AAGATGAACAATTCCAAGCAGATGGAGTGACAGTCCTAAATAGTGTTGAA
CAAGTTATAAAAATGGTTTCAGGAACATAATAAGACCTTATTTATTGTAGGT
GGTGCAAGTATTTATAAAGCATTCTGCCTTATTGTGAAGCAATCATAAAA
ACTAAAGTTCATGGAAAATTCAAAGGTGATACCTATTTTCCTGATGTTAAT
CTATCTGAGTTT

40 MTKQIIAIWAEDEDHLIGVNGGLPWRLPKELHHFKETTMGQALLMGRKTFDG
MNRRLVPGRETIILTKDEQFQADGVTVLNSVEQVIKWFQEHNKTLFIVGGASI
YKAFLPYCEAIKTKVHGKFKGDTYFPDVLNLFSEF

Sequence description:

- 5 A] Length: 417 bp - 139 aa (partial sequence)
 B] ATG start codon is preceded by a Shine-
 Dalgarno sequence. No leader peptide sequence

10 ID-142

Clone 2-c78

15 TTGTGGCCAAACTGTGCCCCGCTTATTAATAGCACTTTGTTCCACCATTGAA
 GATATCTTAACATCAGGTGCTCATAGCAACCCTATTTTAATGGGGGTTATA
 CTTGGCGGGACAATTGTAGTAGTGGCGACAGCACCACCTTTCTTCTATGGCA
 TTGACAGCTATGCTAGGATTAACCGGAATGCCTATGGCTATAGGAGCCTTG
20 TCTGTCTTTGGTTCGTCATTTATGAATGGTGTACTTTTCCATAAATTAAAAC
 TTGGAAGTCGTAAAGATAATATAGCTTTTGCTGTTGAGCCTCTAACTCAAG
 CTGACGTGACTTCAGCTAACCCTATTCCAATCTATGTCACTAATTTTGTTGG
 TGGTGCAGCTTGTGGTATTTTAATTGCCTTGATGAAATTAGTTAATGATACT
 CCTGGAACAGCGACACCAATTGCAGGATTTGCTGTCATGTTTGCCTATAAC
25 CCAATGATAAAAGTACTAATAACCGCTCTAGGTTGTATTATCCTATCTTTA
 CTAGCAGGCTATTTTGGAGGCATTGTTTTT

 MWPNCAPLINSTLFTIEDILTSGAHSNPILMGVILGGTIVVVATAPLSSMALTA
 MLGLTGMPMAIGALSVFGSSFMNGVLFHKLKLGSRKDNIAFAVEPLTQADV
30 SANPIPIYVTNFBVGGAAACGILIALMKLVNDTPGTATPIAGFAVMFAYNPMIKVL
 ITALGCIILSLLAGYFGGIVF

Sequence description:

- 35 A] Length: 540 bp - 180 aa (partial sequence)
 B] N- and C-termini have yet to be elucidated

40

ID-143

ID-145

5

Clone 3-86

10 ATGTCATATTTTAGAAATTACTGGTATCGTTTTGGAGCAATTTTATTTATTA
TTTTAGCAGTAATATTGCTTGTTTTTAGACCTGACTGGTCAATGCTTCACTA
TCTATTGTATTTTTACTTTATGGCACTTCTAGCGCATCAATTTGAAGAATAT
CAGTTTCCCGGTGGGGCATCACCTATCATTAACTATGTTGTTTATGATGAA
GAAGAGCTGATGGATTGTTTTCCAGGCAATACTCAGTCTATTATGTTGGTT
AATACTATTGCTTGCTTGCTTTACATTGCTAGTATTGCTTTTCCTCAAGCTT
15 ATTGGCTTGGATTAGGAGTCATGTTCTTTAGTCTAACGCAGCTCTTGGGTC
ATGGTTTTTCAGATGAATATTAAACTTAAACTTGGTATAATCCTGGTCTAG
CAACGACAGTATTTCTCCTAGTACCAATAGCTTGCGCATACATCTATCAAG
CTAGTGCAGAAGGAATGCTCACTTGGGGAGATTGGCTAGGTGGTTTTATCA
TGTTGATTGTCTGTGTACTAACTAGCATTATTGCACCTGTACAGCTATTGAA
20 GGATAAGGAGACCAATTATATTATTAGTCCTTGGCAAATGGACCGTTTTCA
TAAGGTCGTTAATTTTGTAAGGATAAAAAAATAA

25 MSYFRNYWYRFGAILFIILAVILLVFRPDWSMLHYLLYFYFMALLAHQFEEYQ
FPGGASPIINYVVYDEEELMDCFPGNTQSIMLVNTIAWLLYIASIAFPQAYWLG
LGVMFFSLTQLLGHGFQMNIKLKTWYNPGLATTVFLVPIACAYIYQASAEG
MLTWGDWLGGFIMLIVCVLTSHAPVQLLKDKETNYIISPWQMDRFHKVVNFV
RIKK*

30 Sequence description:

35 A] Length: 651 bp - 219 aa (full length gene)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

40 ID-146

Clone 3-c88

Clone 2-c80

5 ATGTTTTTAAAGTATAATGGCAGGTGTCATAGCATTGTCTGACAGTTATT
 GCCATTCCACGCTTCATTAAGTTTTACCAATTGAAGAAAATTGGCGGGCAA
 CAAATGCATGAAGATGTCAAACAACATCTAGCCAAAGCAGGTACGCCGAC
 AATGGGAGGAACGGTATTTT

10 MFLSIMAGVIAFVLTVIAIPRFIKFYQLKKIGGQQMHEDVKQHLAKAGTPTMG
 GTVF

Sequence description:

15 A] Length: 172 bp - 57 aa (partial sequence)
 B] Shine Dalgarno sequence precedes 'ATG' start
 codon. Possesses a potential leader peptide
 sequence.

20

ID-144

25

Clone 3-83

30 ATGAAACCATATTTATCTTTTATTGGTAGAACGTTATTATACTTCGGTATTT
 TATTGTTACTAATTTACTTTTTTGCATACCTTGGTCGCGGACAAGGCAGTTT
 TATTTATAA

MKPYLSFIGRTLTYFGILLLLIYFFAYLGRGQGSFIY

35

Sequence description:

40 A] Length: 113 bp - 37 aa (partial sequence)
 B] Putative ATG start codon is preceded by a
 typical Shine-Dalgarno sequence. Possesses a
 potential leader peptide sequence.
 This orf is not in frame with nuc

ATGCCACTTACAGCACTTGAAATTAAAGATAAAACATTTTCATCAAAATTT
CGCGGTTATAGCGAAGAAGAAGTT

5 MPLTALEIKDKTFSSKFRGYSEEEV

Sequence description:

10

A] Length: 75 bp - 25 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No leader
peptide

15

ID-147

20 Clone 3-90

ATGTCACCTTTTTCAAGAAAAAATTGCTTACAATTGCGCTAAAAAGGAAGCG
CTTTATAAAGAGAGTTTAGGACGCTACGCCTTGAGATCAATGCTAGCAGG
25 GGCTTATTTGACAATGAGTACTGCTGCCGGTATCGTCGCAGCTGATACTAT
TGGTAAAATTTCTCCTGCTCTATCAGGTTTTGTATTTGCTTTTCATCTTTAGTT
TTGGACTTATTTATGTTTTAATATTTAATGGTGAATTGGCGACATCTAATAT
GCTTTATCTCACTGCAGGAGCCTATAATAAAAATATCTCTTGGAAAAAAGC
CATAACAATTTTAATTTATTGTACTTTTTTCAACCTCGTTGGTGCTTGTATA
30 TTAGCTTGGTTGTTTAA

MSLFQEKIAYNCAKKEALYKESLGRYALRSMLAGAYLTMSTAAGIVAADTIG
KISPALSGFVFAFIFSFLIYVLIFNGELATSNMLYLTAGAYNKNISWKKAITILI
YCTFFNLVGACILAWLF

35

Sequence description

40

A] Length: 406 bp - 125 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possible
leader peptide

ID-148

5

Clone 3-92

10 AAGTTACAAGCGACTGAAGTTAAGAGCGTTCCGGTAGCACAACCAGCTTC
AACAACAAATGCAGTAGCTGCACATCCTGAAAATGCAGGGCTCCAACCTC
ATGTTGCAGCTTATAAAGAAAAAGTAGCGTCAACTTATGGAGTTAATGAA
TTCAGTACATACCGTGCGGGAGATCCAGGTGATCATGGTAAAGGTTTAGC
AGTTGACTTTATTGTAGGTAAAAACCAAGCACTTGGTAATGAAGTTGCACA
15 GACTCTACACAAAATATGGCAGCAAATAACATTTTCATATGTTATCTGGCA
ACAAAAGTTTTATTCAAATACAAATAGTATTTATGGACCTGCTAATACTTG
GAATGCAATGCCAGATCGTGGTGGCGTTACTGCCAACCCTATGACCACGT
TCACGTATCATTTAA

20 KLQATEVKSPVAQPASTTNAVAAHPENAGLQPHVAAAYKEKVASTYGVNEF
STYRAGDPGDHGKGLAVDFIVGKNQALGNEVAQYSTQNMAANNISYVIWQQ
KFYSNTNSIYGPAANTWNAMPDRGGVTANHYDHHVHVSF

Sequence description

25

A] Length: 419 bp - 139 aa (partial sequence)
B] N- and C-termini have yet to be determined

30

ID-149

35

Clone 3-94

40 ATGATTCCAGTAGTTATTGAACAAACAAGTCGTGGTGAACGTTCTTATGAT
ATTTACTCACGTCTTTTAAAAGATCGTATTATTATGTTGACAGGCCAAGTT
GAGGATAATATGGCCAATAGTATCATTGCACAGTTATTGTTTCTCGATGCA
CAAGATAATACAAAGGATATTTACCTTTATGTCAATACACCAGGTGGTTCA
GTATCGGCTGGACTTGCTATTGTGGACCATGAACTTCATTAAATCGGAC
GTACAGACGATTGTTATGGGGATGGCTGCTTCGATGGGAACCATATTGCT
TCAAGTGGTGCTAAAGGAAAACGTTTTATGTTACCGAATGCAGAATATATG

ATCCACCAACCAATGGGCGGAACAGGCGGAGGTACACAGCAATCTGATAT
 GGCTATCGCTGCTGAGCATCTTTTAAAAACGCGTCATACTTTAGAAAAAAT
 CTTAGCTGATAATTCTGGTCAATCTATTGAAAAAGTCCATGATGATGCAGA
 GCGTGATCGTTGGATGAGTGCTCAAGAACACTTGATTATGGCTTTATTGAT
 5 GCTATTATGGAAAATAATAATTTACAATAATAGATTTAAAAGAGTTGAGTT
 TACCAACTCTTTTTTTTATTTGTTGGAATTATGTTATAATCTTAGTAATTACA
 GATATGACGCAGAAAGGAAAAAATTATTGA

MIPVVIEQTSRGERSYDIYSRLLKDRIIMLTGQVEDNMANSIIAQLLFLDAQDN
 10 TKDIYLYVNTPGGSVSAGLAIVDTMNFIKSDVQTIVMGMAASMGTHASSGAK
 GKRFMLPNAEYMIHQPMGGTGGGTQQSDMAIAAEHLLKTRHTLEKILADNSG
 QSIEKVHDDAERDRWMSAQEHLIMALLMLLWKIIYNNRFRKVEFTNSFFICW
 NYVILVITDMTQKGKNY*

15

Sequence description

20

A] Length: 693 bp - 231 aa (full length gene)
 B] Putative ATG start codon is preceded by a
 typical Shine-Dalgarno sequence. No leader
 peptide. Significantly, it would appear to have a
 very hydrophobic C-terminus.

25

ID-150

30

Clone 2-c86

35

40

ATGAAACCAAAAATTATTGGTGTACTTGGTCTAGGAATATTTGGACAAACA
 CTCGCACAAGAATAAGTAACCTTTGAACAAGATGTTATTGCTATTGACAGC
 AATCCTGAAAATGTACAAGCTGTCGCCGAAGT
 TGTTACAAAAGCAGCTATCGGAGACATTACTGATTTAGCTTTCCTAAAACA
 CATCGGGATCAGTGACTGTGATACTGTTATTATTGCTACAGGAAACAGTTT
 AGAGAGCTCAGTATTGGCCGTAATGCACTGTAAAAAGTTAGGCGTCCCAC
 AAGTTATTGCTAAAGCTCGAAACCTTGATACGAAGAAGTACTTTATGAAA
 TTGGTGCTGATTTGGTTATCTCTCCGGAGCGAGAATCTGGGCAAAATGTTG
 40 CTGCAAACCTCATGAGAAATAAAATTACAGATGTCTTCCAGATTGAATCTG
 ATATTTCTGTCATTGAATTT

MKPKIIGVLGLGIFGQTLAQELSNFEQDVIAIDSNPENVQAVAEVVTKAAIGDI
 TDLAFLKHIGISDCDTVIIATGNSLE

SSVLAVMHCKKLGVPQVIKARNLVYEEVLYEIGADLVISPERESGQNVAAN
LMRNKITDVFQIESDISVIEF

5 Sequence description:

- 10 A] Length: 459 bp - 153 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.
This orf is not in frame with nuc

15 ID-151

20 Clone 2-c88

25 GTGCGTTATAGTAAAGAGATTATTCAGTTAGCTATACCAGCTATGATTGAA
AATATCTTACAAATGCTCATGGGAGTAGTTGATAATTATCTAGTGGCTCAG
TTAGGTGTTGTAGCAGTATCAGGTGTTTCAGTTGCTAATAATAATTACT
ATTTATCAAGCTATTTTTATAGCTTTAGGGGCGAGTATAGCAAGTCTATTG
GCCAAGTCGTTAGCAGGTAGTGAGAAGGATGATGCAATTTTCAGTATGTTCT
CAAGCCATTTTTCTAACATCACTGATAGGGGCAGTATTAGGAATTATCTCG
ATTGTTTTTGGACAAACTTTCTTT

30 MRYSKEIIQLAIPAMIENILQMLMGVVDNYLVAQLGVVAVSGVSVANNIITIY
QAIFIALGASIASLLAKSLAGSEKDDAISVCSQAIFLTSLIGAVLGIISIVFGQTFF

35 Sequence description

- 40 A] Length: 330 bp - 110 aa (partial sequence)
B] Putative GTG start codon is preceded by a
typical Shine-Dalgarno sequence. May have a
leader peptide

ID-152

Clone 2-c92

5

TTGATTAACAAGTATTCGTGCTTTTTGAAGAGGATTCTCCATAATAATACT
CCTTTAATAGTTATCGTGAGAAGTATTTTAAAGAAAAACCGCCAAGGTAG
AGCGACATTTCTGCCTTTAACTACAATAAAACCAAGAGAATTAGCACAAAC
ATTATCTCTCAAAATTACAAAGTTCTCAAGGGTTTTTAGGAATAGCTAGTG
AATTGGTAACCTATGATCAACGCTTGTCAAACATTTTT

10

MINKYSCFLKRILHNNTPLIVIVRSILKKNRQGRATFLPLTTIKPRELAQHYSK
LQSSQGFLGIASELVTYDQRLSNIF

15

Sequence description

A] Length: 240 bp - 80 aa (partial sequence)

B] No obvious Shine Dalgarno sequence precedes the Putative TTG start
codon

20

ID-153

25

Clone 2-c94

30

TTGTTGACTCACAAAAATATATTATTAACCATTATATTTGGATTATTTATGA
TTATATTATCAGCATGTGGTATGTCTAATAAGGAAATGGCTGGTATTGATA
ATTGGGAACATTATCAAAAGGAAAAGAAAATTACTATTGGATTTGATAAT
ACTTTTGTTCCCTATGGGATTTGAAAGTCGTTCTGGTGACTATACCGGCTTTG
ATATTGATTTAGCTAATGCTGTTTTTAAAGAATACGGTATTTTCAGTGAAAT
GGCAGCCTATTAACCTGGGATATGAAAGAACTGAACCTAATAATGGTAAT
ATAGACCTTATTTGGAATGGTTATTCAAAAACGGCAGAACGTGCTAAAAA
AGTCGCTTTTACAAACCCATATATGAATAATCATCAAGTAATTGTTACTAA
AACTTCATCACATATTAATAGTATTAAGGATATGAAGGGGAAAAAACTAG
GAGCCCAGTCGGGTTTCATCTGGTTTTGATGCTTTTAACGCTAAACCTGATA
TTTTAAAAAAGTTTGTAAGGAAAAGAAGCAGTTCAATACGATACTTTC
ACTCAGGCTTTGATTGATTTAAAAAATAACCGTATTGATGGTCTTTTGATT
GATGAAGTTTATGCTAACTATTATTTAAAGCAAGAAGGAA

40

MLTHKNILLTIIFGLFMIILSACGMSNKEMAGIDNWEHYQKEKKITIGFDNTFV
PMGFESRSGDYTGFDIDLANAVFKEYGISVKWQPINWDMKETELNNGNIDLI
WNGYSKTAERAKKVAFTNPYMNNHQVIVTKTSSHINSIKDMKGKKLGAQSG
SSGFDAFNAKPDILKKFVKGKEAVQYDTFTQALIDLKNNRIDGLLIDEVYANY
5 YLKQEG

Sequence description

10

A] Length: 649 bp - 216 aa (partial sequence)
B] TTG start codon is preceded by a possible
typical Shine-Dalgarno sequence. Has a
leader peptide

15

ID-154

20

Clone 2-c100

25 ATGAAAATTTGGAAAAAATAACCTTAATGTTTTCTGCAATTATTTTAACA
ACAGTAATTGCATTGGGAGTCTATGTTGCCTCAGCTTATAATTTTTCGACTA
ATGAATTGTCTAAGACTTTT

MKIWKKITLMFSAIILTTVIALGVYVASAYNFSTNELSKTF

30

Sequence description

35 A] Length: 123 bp - 41 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide

40

ID-155

Clone 2-c1

ATGAAAAACAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
5 ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
GGAATACCAGGCAGAACAGAAATTTTAAGTCATACTTTAAATATATATCAG
ATAAAAATAACTATTTAGATAATATAAAAGTTTATTACTTTTCTATAAGTA
TTTCTAAAGATGTACAAGATAAAGTCAGTGAAACAACAACCTTGTTTCATATA
GACTAGAAAAGCAAAAGAATCAAGAGTTCATTGGTAATTTTGAACATGAA
GTTAGTGAATCTAGTCAATATTCAACCGAAGTTAAAAATCAAATACAGTAT
10 CCAATCCAGTATAAAGATAATTCAATTCGTTTTACTGAAAAAACACCGTCA
GAACGTTATGATGAGTTTGTGTTTTAGTTCATTGATTCTTCATTATTA
AATATAAAATATATGATTACTTACTAAAACATCCCGAAACTGAATTA
GGTGTTCCTATAAGATTCCTATAAATTCTGAAATTGTAGCCCCTTTTATAA
ATCAATTAAATATAAAAAATCCTAAAAAATCATCTATTTTCGGTTACAAAAA
15 CGGAAAGTAAAGAATATTATTATACAATCAGTATTGATACTGATTCTGAGA
TATATTCTATATTCGAAGGTATTCAT

MKKQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYFKYISDKN
20 NYLDNIKVYYFSISISKDVQDKVSETTTCYSRLEKQKNQEFIGNFEHEVSESSQ
YSTEVKNQIQYPIQYKDNSIRFTEKTPSERYDEFVFSSFDSSLLKKYKIYDYLLK
HPETELKGVSYKIPINSEIVAPFINQLNIKNPKKSSISVTKTESKEYYYTISIDTDS
EIYSIFEGIH

25 Sequence description

30 A] Length: 687 bp - 229 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide. C-terminus has yet to be
verified

35 ID-156

Clone 2-c5

40 ATGACATTTGACACCATTGATCAATTAGCGGTTAATACAGTCCGCACGCTT
TCTATTGATGCTATCCAAGCAGCAAATTCTGGGCACCCAGGTCTTCCTATG
GGAGCTGCGCCTATGGCTTATGTGCTTTGGAATAAATTCTTAAATGTAAAC
CCAAAAACAAGTCGCAATTGGACAAACCGTGACCGTTTTGTACTTTCAGCT

GGGCATGGTTCAGCTCTTCTTTATAGCCTACTTCATTTAGCTGGCTATGATT
TATCAATTGATGATTT

5 MTFDTIDQLAVNTVRTLSIDAIQAANS~~GHPGLPMGAAPMAYVLWNKFLNVNP~~
KTSRNWTNRDRFVLSAGHGSALLYSLHLAGYDLSIDD

Sequence description

10

A] Length: 272 bp - 90 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. No obvious
leader peptide

15

ID-157

20

Clone 2-c8

25 ATGAGAACACTATTTAGAATGATATTTGCTATTCCAAAGTTTATCTTTAGA
TTGATTTGGAATATCATTTGGGGAATATTCAAGACAGTTCTTGTTATTGCG
ATTATTTTATTTGGCTTGTATTACTATGCGAATCACAGTCAATCAGAATTTG
CTAATCAACTTAGTGACATTATTCAGACAGGAAAAACATTTTT

30 MRTLFRMIFAIPKFIFRLIWNIIWGIFKTVLVIAIILFGLYYYANHSQSEFANQLS
DIIQTGKTF

Sequence description

35

A] Length: 197 bp - 65 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

40

ID-158

Clone 2-c9

5

ATGTCAAAAAAAAAATAATATTAGGAATTTTATCTCTTTTATCTGTCGTTACTT
TGGTGGCGTGTGGTTCATCAGACAAACAGCTACAAGATAAAGTTGAGAAA
AAAGGGAAGTTAGTTTTAGCGGTGAGTCCAGATTATGCTCCCTTTGAGTTT

10

MSKKIILGILSLLSVVTLVACGSSDKQLQDKVEKKGLVLAVSPDYAPFEF

Sequence description

15

A] Length: 153 bp - 51 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide (not in frame with nuc)

20

ID-159

25

Clone 2-c10

ATGAAAAATCAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
30 GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTT

MKNQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYF

35

Sequence description

40

A] Length: 139 bp - 46 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

ID-160

Clone 2-c11

5

ATGATTGGAAAATTATATTATAGCTATAGAAAGTCACGCTTATTAAGAAGT
ATTTTATGGCTTATTTTAATTGTTGGTGTATATATGTTAGGACAACGTGTTT
TATTATCCACTGTTTCCTTTATCACATCAAGAGATAAACTAGCAGTAGATC
10 AACATTTACTCAATAACTTTTCAGCAGTAAGTGGTGGGAGTTTTTAATAAAT
TAAATGTTTTCACACTGGGGTTGAGTCCATGGATGTCAAGTATGATTATTT
GGAGATTTCGTTTCCTTATTTTCGTGGGCAAAAAATGCAACGAAGCGAAAA
GCAGAAGTAGCTCAATATACTTTAATGCTTACTATCTCAGTTATACAAGCA
TATGGTGTTCAGGAAATCAATTTATAAAAAGCTCTTTATTAGGTTCTTATA
15 GTGATATTGTTTTT

MIGKLYYSYRKSRLRLSILWLILIVGVYMLGQRVLLSTVPLSHQEIKLAVDQHL
LNNFSAVSGGSFNKLVFTLGLSPWMSSMIIWRFVSLFSWAKNATKRKA EVA
QYTLMLTISVIQAYGVSGNQFIKSSLLGSYSDIVF
20

Sequence description

25

A] Length: 423 bp - 141 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

30

ID-161

35

Clone 2-c13

ATGAAAGGTCTATTGGATTTTTTTAGTTAATATTGCCAGAACGCCAGCTATT
TtagTCGCCTTGATAGCCATTATCGGTTTtagTACTGCAGAAAAAAGGTGTT
40 CCTGATATTGTAAAAGGTGGAATAAAAAACATTTGTTGGCTTCTTAGTGGTT
TCTGAAGGTGCAGGGATAGTCCAAAATTCCTTGAATCCATTTGGAAAAATG
TTTGAACATGCTTTTCATTTGGTGGGGGTAGTTCCTAATAATGAAGCCATT
GTAGCAGTAGCTCTTACGAAGTATGGCTCAGCAACTGCTTTGATTATGTTA
GCGGGAATGATTTTTAATATTTTAATTGCTCGTTTTACAAAA

MKGLLDFLVNIARTPAILVALIAIIGLVLQKKGVDPDIVKGGIKTFVGFLVVSEG
AGIVQNSLNPFGKMFEHAFHLVGVVPNNEAIVAVALTKEYGSATALIMLAGMI
FNILIARFTK

5

Sequence description

- 10 A] Length: 348 bp - 116 aa (partial sequence)
 B] ATG start codon is preceded by a potential
 Shine-Dalgarno sequence. Possible leader
 peptide

15

ID-162

Clone 2-c21

20

TTGGTTGGTAAGCCCCAATTACTATTTTTAGATGAACCTACTTCCGGAATG
GATACTTCCACACGTCAACGATTTTGGAAGCTGGTTGCGACACTAAAAAA
AGAAGGTGACACAATTGTCTATTCTAGTCATTATATCGAAGAGGTTAGAAC
25 ATACAGCTGATAGGATTTTAGTACTTCATAAAGGAAAGTTATTACGCGATA
 CAACCCCCTTTGCCATGAAGCAAGAAAAAACCGAAAAGTTATTCACCGTT
 CCGCTTAGTTATCAAAAATTATTACCTACCTATTTGATTACAGAGTGTGAA
 GCCAAGAGTGATAGTATAACGTTTGTTACTGGGGAGGCTGAAACTGTATG
 GAAAATACTGGCAGATAATGGTTGTCCTATTGAAGCTATTGAGATGACCA
30 ATAGAACTTTGTAAATCGTATTTTTGAGACTACTAAGGAGGTAAACATG
 AGAATCTTTA

35

MVGKPQLLFLDEPTSGMDTSTRQRFWKLVALTKKEGDTIVYSSHYIEEVEHTA
DRILVLHKGKLLRDTPFAMKQEKTEKLFTVPLSYQKLLPTYLITECEAKSDSI
TFVTGEAETVWKILADNGCPIEAIEMTNRTLLNRIFETTKEVKHENL

Sequence description

40

- A] Length: 462 bp - 155 aa (partial sequence)
B] B] Putative TTG start codon is not preceded by
an obvious Shine-Dalgarno sequence. No obvious
leader peptide. N- and C- termini require further

examination.

ID-163

5

Clone 2-c25

10

TTGAAAAAATCCAAGAGAAGCCGTAAGGCAGTGACAACAAGTGGTGAGA
AGACTTTACTTGAGGATTTGGCAAAAATGAATTCCTAGACGAAGTCATTA
ATGTTATGGTTTTATATACCTTGAATAAGACAAAATCTGCTAACTTAAATA
AGGCCTATATCATGAAAGTTGCTAATGATTTTGCCTTTCAGAATGTTATGA
CGGCCGAAGATGCTGTGCTTAAAATTCGTGATTTTTCAGATCAAAAAGTAA
GGACTAAAACAGAAACGAAGAAGAAACAATCGAATGTTCTGAATGGAGT
15 AATCCTGATTATAAAGATGAGGTTAGCCCAGAAAAAGAAATTGAATTAGA
ACAGTTT

20

MKKSKRSRKAVTTSGEKTLLEDLAKMNFLDEVINVMVLYTLNKTKSANLNK
AYIMKVANDFAFQNVMTAEDAVLKIRDFSDQKVRTKTETKKKQSNVPEWSN
PDYKDEVSPKEIELEQF

Sequence description

25

A] Length:360 bp - 120 aa (partial sequence)
B] N- and C- termini require verification.

30

ID-164

35

Clone 2-c28

40

ATGACGAATCATATTACTAAACTGATAGAAAATAGCGGAAAAAAATTGAC
AGAAATTAGCGAAGCTACAGATATAGCCTATCCTACACTTTCTGGATACAA
TCAAGGAATCCGCAAACCTAAAAAAGATAATGCTGAAAAATTGGCAAAAT
ACTTTAATGTTTCCGTCGCTTACATTATGGGACTTGATAGCAACCCACATG
CTCCATCAAATCTT

MTNHITKLIENSGKKLTEISEATDIAYPTLSGYNQGIRKPKKDNAEKLAKYFNV
SVAYIMGLDSNPHAPSNL

Sequence description

5

A] Length:218 bp - 72 aa (partial sequence)
B] ATG start codon is preceded by an
obvious Shine Dalgarno sequence. No obvious
leader peptide.

10

ID-165

15

Clone 2-c29

20

TTGATGAAAAGGAATAAACATTTACCGTTAACAGAACTACCTATTATATT
TTATTAGCTTTGTTTGAGGAAGCGCATGGCTATGCTATTATGAAAAAAGTT
GAAGAAATGAGTGGCGGTGATGTTAGAATAGCCGCAGGGACAATGTACGG
TGCCATTGAAAATTTACTTAAACAAAAATGGATAAAGTCTATCTCAAGTGA
CGATAGAAGAAGAAAAGTTTATATTATTACTGAGACAGGAAAAGAAATAG
TAGAACTTGAAACGAATCGATTAAAGAAAGTTACTTAATACTGCTAATCAGT
TGGGTTTTGGAGGAGATGGTTATGATAAAGTTT

25

MMKRNKHLPLTETTYIILLALFEEAHGYAIMKKVEEMSGGDVRIAAGTMYG
AIENLLKQKWIKSISSDDRRRKVYIITETGKEIVELETNRLRKLLNTANQLGFG
GDGYDKV

30

Sequence description

35

A] Length:337 bp - 112 aa (partial sequence)
B] TTG start codon is preceded by an
obvious Shine Dalgarno sequence. Actual start
codon may ATG that comes immediately after the
TTG. Potential leader peptide.

40

ID-166

Clone 2-c35

5 CCCATTACTGGTGAGTTAATAGCTGAGAAATTAGGAGTACCAAGAGCAGC
ACTAAGGTCTGATTTGCGGGTTTTAAGTATGCTAGGTATCATAGATGCAAA
ACCTAAGGTTGGTTATTTTTATTTAGGACAGTATCATGCTTCAATAGGGAC
AAGTCATTTTGAAAAGATGACAGTTTCAGAAATTATGGGGATCCTTCTGAC
AGTTCATCAAAAAGATTCAGTTTATGATGTTATTGTACATATTTTTATGGA
10 AGATGCTGGTTGTGCTTTTATCTTGGATGATGATGATTTTCTCTGTGGAGTC
GTGTCACGTAAAGATTTACTAAAAACCAGTATTGGCGGAGGAGATCTTTCT
AAAATGCCAATAGGAATGGTGATGACACGTATGCCACACGTGACAACCTGT
TTTAGAAAATGAAAGTCTTTTTGCGGCAGCTGATAAATTAGTGAGCAGAA
AAGTGGATAGTCTCCCTGTCGTTTCGTCATGATAAGCAATATCCCGAAAAAT
15 TTA

PITGELIAEKLGVPRALRSDLRVLSMLGIIDAKPKVGYFYLGQYHASIGTSHF
EKMTVSEIMGILLTVHQKDSVYDVIVHIFMEDAGCAFILDDDDFLCGVVSRKD
LLKTSIGGGDLKMPIGMVMTRMPHVTTVLENESLFAAADKLVSARKVDSL PV
20 VRHDKQYPEKF

Sequence description

25 A] Length:511 bp - 170 aa (partial sequence)
B] N- and C-termini to be determined

30 ID-167

Clone 2-44

35 TTGGAAGTCATCATGCAATTTATTTATAGTATTATTGGTATTTTATTGGTAT
TAGGAATTGTGTATGCAATTTCTTTCAATCGTAAGAGTGTTTCTCTAAGTTT
AATTGGAAAAGCTCTTATCGTTCAATTCATTATTGCGCTAATCTTAGTACGT
40 ATCCCACTAGGCCAACAAAGTTGTTAGTGTTGTTTCAACTGGAGTTACTAAA
GTAATCAACTGTGGTCAAGCTGGTTT

MEVIMQFIYSIIGILLVLGIVYAISFNKRKSVLSLIGKALIVQFIILVRIPLGQQ
VVSVVSTGVTKVINCGQAG

Sequence description

5

A] Length:233 bp - 77 aa (partial sequence)
B] TTG start codon is preceded by a
possible Shine Dalgarno sequence. Actual start
codon may occur further downstream. Potential
leader peptide.

10

ID-168

Clone 2-46

15

CAACCTAATAAAGCTTTAGAAAGTGATGAGATTGATATTAATGCTTTCCAG
CATTATAATTACTTAACCAATTGGAATAAAGCAAATAAGACCAATCTTGTT
TCCGTTGCTGAGACATACTTTACTTCCTTTAGATTATACTCTGGTACTAAGA
20 ACGGTAAAGGTAAATACCAAACAGTTTCTGAAATTCCAAATAAAGCAACT
ATTACTATCCCAAACGATGCAGTTAACGAAAGTCGCTCTCTCTACTTGTTA
CAATCAGCAGGCTTGCTAAAATTGAAAGTATCAGGTGATACATTAGCAAC
AATGTCAGATGTTGTTTCCAATCCTAAATCTTTAGATT

25

QPNKALESDEIDINAFQHYNLTNWNKANKTNLVSVAETYFTSFRLYSGTKN
GKGKYQTVSEIPNKATITIPNDVNESRSLYLLQSAGLLKLKVSGDTLATMSD
VVSNPKSLD

30

Sequence description

A] Length:344 bp - 114 aa (partial sequence)
B] N- and C- termini require verification

35

ID-169

40

Clone 2-47

ATGAAATGTATAATAAATAATATAAAATAAAATAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAACATTTTGAAATTAAATAAAATCATATTATCTAC

TGCAGCTCTTACTGCTCTCTTTTTAGGATATAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTTAACTGAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

5

MKCIINNINKIKMIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

10

Sequence description

15

A] Length:264 bp - 88 aa (partial sequence)
B] There is a Shine-Dalgarno sequence upstream
of this sequence. Potential leader peptide
sequence

20

ID-169

Clone 2-47

25

ATGAAATGTATAATAAATAATATAAAATAAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAACCTATTTTGAAATTAAATAAAATCATATTATCTAC
TGCAGCTCTTACTGCTCTCTTTTTAGGATATAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTTAACTGAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

30

MKCIINNINKIKMIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNNNQIELQERIKNLNIPF

35

Sequence description

40

A] Length:264 bp - 88 aa (partial sequence)
B] There is a Shine-Dalgarno sequence upstream
of this sequence. Potential leader peptide
sequence

ID-170

5 Clone RS-58b

TTGGGTGATTATTATGGTAAGAAATATTTTGGTGAGGCAGCTAAAAAAGA
CGTCGAACATATGGCTAAGAAAATCATTAATGTCTATAAAACACGGTTAA
10 AAAACAACACTTGGTTATC
AGAAAATACAAAAGCAATGGCCATTAAGAAACTTGATAACATGAGATTAA
TGATTGGCTATCCAGAAGATTATCCTGATCTTTATCGTCAGTACCAATTTG
ATAGTAAAGCAAGCTTCTTTGAAAACAATGATAACTACAGAAAATTATCG
AACAAGAAAACATTTGAAGAATTTAACCAGTCTAATCAACGTGAACATTG
15 GCAAATGAGTGCCAATGCTGTAAATGCTTATAATGATCCTAATACCAATTC
CATAGTCTTTCCAGCAGCGATTTTTCAATCACCCTGTACGATAAACTAA
AACAGTTAGTCAAAATTATGGAGCTATCGGAGCAATTATTGGTCATGAAAT
TTCACACTCATTTGATATTAATGGTATGAAATATGACGAGAAAGGGAATCT
TCACGATTGGTGGACTAAAGAAGATTTAAATCATTATAAGAAATCAACAC
20 AAGCTATGATTGACCAATGGGATGGCCTTAAAGCAGATGGCGGTAAAGTT
GATGGTAAATTAACCTTTAGCAGAAAATATTGCAGATAATGGTGGTGTATG
GCATCTCTAGAAGCTCTTAAGACTGAAAAAATCCAACTATAAAGAATTTT
TTGAATCATGGGCAAGTATTTGGCGTCAAAAAGCAACCAAAGAACAAGT
AAGTCCTCAATTCAGTCAGATGTTTCATGCACCATATGAATTGA >
25 GAGCTAACATCCCAGTACGTAATTTCCAAGAATTTTATGATGCCTTTGGTG
TTAAAAAAGGCGATTCAATGTATCTAAAACCAGAAAAACGTTTGACACTTT
GGTAA

30 MGDYYGKKYFGGEAAKKDVEHMAKKIINVYKTRLKNNTWLSENTKAMAIKK
LDNMRLMIGYPDYPDLRQYQFDSKASFFENNDNYRKLSNKKTFEEFNQSNQ
REHWQMSANAVNAYNDPNTNSIVFPAAIFQSPLYDKTKTVSQNYGAIGAIIGH
EISHSFDINGMKYDEKGNLHDWWTKEDLNHYKKSTQAMIDQWDGLKADGG
KVDGKLTAEINIADNGGVMASLEALKTEKIQTIKNFLNHGQVFGVKKQPKNK
35 VSPQFSQMFMHHMN*

Sequence description:

40 A] Length: 819 bp - 272 aa (full length gene)
(107 bp of additional DNA sequence (> onwards) is
also included. While not in-frame with the
described orf, it also shares strong homology
with the neutral peptidases.

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-89 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-89 gene sequence. ID-89 and ID-170 together show homology over their combined entire length with the neutral endopeptidases from *Lactococcus* and *Lactobacillus*. Possesses TTG (possible ATG start codon located 13 bp further downstream) start codon with no obvious signal peptide. Shine Dalgarno sequence not immediately obvious. Possibly located further downstream

10 ID-171

Clone 2-18/22b (Mod2)

15

ATGACCATGATTACGCCAAGCTTCATTAAGGTATCTCTAGATGAAACAAAT
CGTATGATGCGTATGATATCAGATTTATTAAGTTTATCGCGCATTGATAAT
GAAGTAACGCATTTAGATGTTGAAATGACGAATTTTACAGCTTTCATGACC
TCAATTTTGAATCGATTTGATCAGATTAGAAATCAAAAAACAGTCACAGG
20 AAAAGTTTATGAAATTGTCAGAGATTATCCTCTTAAGTCAATTTGGGTGGA
AATTGATACAGATAAGATGACTCAAGTGATTGATAACATTTTAAATAATGC
AGTCAAGTATTCACCAGATGGTGGTAAGATTACAGTTAATCTACGCACAAC
TAAAACGCAGATGATTTTATCAATATCAGACCAAGGCTTAGGTATTCCCAA
AAAAGATTTACCTCTCATTTTTGATCGTTTTTATCGTGTTGATAAGGCGAGA
25 AGTCGTCAACAGGGTGGGACTGGACTTGGTTTGTCAATTGCAAAAGAAAT
TGTTAAGCAGCATAAGGGATTTATTTGGGCTAAGAGTGAGTATGGTAAAG
GGTCTACTTTTACAATCGTCTTGCTTATGATAAAGATGCTGTAACCTTATGA
AGAATGGGAGGACGTTGAAGATTAA

30

MTMITPSFIKVSLDETNRMMRMISDLLSLSRIDNEVTHLDVEMTNFTAFMTSIL
NRFDQIRNQKTVTGKVVEIVRDYPLKSIWVEIDTDKMTQVIDNILNNAVKYSP
DGGKITVNLRTTKTQMILSISDQGLGIPKKDLPLIFDRFYRVDKARSRQQGGTG
35 LGLSIAKEIVKQHKGFIWAKSEYGKGSTFTIVLPYDKDAVTYEEWEDVED*

Sequence description:

- 40 A] Length: 613 bp - 212 aa (full-length gene possibly)
B] Possible Shine Dalgarno sequence present upstream of a ATG start codon. May not have yet determined the N- portion of this gene. No obvious signal peptide.

ID-172

5

Clone 2-54balternate (107b)

10

TTGAAAAAAATTATTACTTCTATTCTATTACTTAGTTGCATTTTTTTTTATGC
CAACCATCTCTGCTGAATCTTTTAATGCTTCCGCTAAACATGCCTTAGCAGT
TGATTTAGATTCAGGAAAAATCTTGTATGAAAAAGATGCTAACAAACCCG
CTGCTATTGCTTCCTTGACTAAAATAATGACCGTTTATATGGTCTATAAAG
AAATTGATAACGGTAACCTCAAGTGGAATACCAAAGTAAATATATCTGAC
TACCCTTATCAACTAACACGCGAATCTGATGCTAGTAATGTTCCTTTAGAA
15 AAAAGGCGCTATACTGTAAACAACCTCGTGGACGCTGCCATGATTTCTAGT
GCTAACAGTGCAGCCATTGCTTTAGCTGAACATATTTTCAGGAACTGAAAGT
AAATTTGTTGATAAAATGACTGCTCAATTGGAAAAGTGGGGAATTCATGAT
AGCCACCTAGTCAATGCTTCTGGCTTAAATAATAGTATGTTAGGCAATCAC
ATTTATCCAAAATCGTCACAAAACGACGAAAATAAAATGAGTGCACGTGA
20 TATTGCTATTGCTGCCTACCATTTGGTCAACGAATATCCTTCCATTCTTAAG
ATTACTAGTAAGTCCGTTGCTAAATTTGATAAAGATATTATGCATTCTTAT
AACTACATGCTACCAGATATGCCTGTCTTTAGACCAGGTATTACAGGTTTG
AAAACCTGGGACAACGGAATTAGCTGGCCAATCTTTTATTGCTACATCTACT
GAAAGTGGAATGAGACTACTCACTGTTATTATGCATGCTGATAAGGCCGAT
25 AAAGACAAATATGCTCGCTTTACAGCAACTAACTCTCTCTTGAACATATATC
ACAAACACCTACGAACCTAACCTTGTATTAGCTAAAGGAGCTGCATATAA
AGGTAAAGAAGCAAGTGTGAGAGACGGAAAAGAACAATCGGTCATCGCT
GTTGCTAAAAACGATTTGAAAGTAGTACAGAAGAAAAATATCACTAAACA
AAATCAGTTAAAAATTAACTTTAAAAAAGAGCTTACTGCTCCTATTACAAA
30 AAAAGAGAACCTAGGGAAAGCTTATTACGTTGACCTTAATAAGGTTGGAA
AAGGCTATCTCATAAAGGAACCTAGCGTTCATTTAGTGGCAAAGATAGT
ATTGAGCGCAGTTTCTTCCTCAAAGTGTGGTGGGAATCATTTTGTGCGCTAC
GTTAACGAAAAACTTTAA

35

MKKIITSILLLSCIFFMPTISAESFNASAKHALAVDLDSGKILYEKDANKPAAIA
SLTKIMTVYMVYKEIDNGNLKWNTKVNISDYPYQLTRESASNVPLEKRRYT
VKQLVDAAMISSANSAAIALAEHISGTESKFVDKMTAQLEKWGIHDSHLVNA
SGLNNSMLGNHIYPKSSQNDENKMSARDIAIAAYHLVNEYPSILKITSKSVAKF
40 DKDIMHSYNYMLPDMPVFRPGITGLKTGTTELAGQSFIATSTESGMRLLTVM
HADKADKDKYARFTATNSLLNYITNTYEPNLVLAKGAAYKGKEASVRDGKE
QSVIAVAKNDLKVVQKKNITKQNQLKINFKKELTAPITKKENLGKAYYVDLN
KVGKGYLIKEPSVHLVAKDSIERSFFLKVWWNHFVRYVNEKL*

40

Sequence description:

- A] Length: 1236 bp - 412 aa (full-length gene sequence possibly)
B] A possible Shine-Dalgarno sequence precedes the putative 'TTG' start codon. (needs further cloning and sequencing to verify N-terminus)

ID-173

Clone 3-60b

ATGACGCTTCGAGAATTAACAATAGAAGAATTTAAAGAACATTCAGGAAA
TTATGATTCACAATCATTTTTACAAACACCTGAGATGGCTAAACTTTTAGA
AAAACGCGGCTATGATGTTAGGTATTTGGGATATCAAGTAGAAAATAAAC
TAGAGATAATCAGTTTATCTTATATTATGCCAGTCACTGGTGGTTTTCAAAT
GAAAATTGATTCAGGACCAGTTCATTCAAATTCTAAGTATCTAAAACAATT
TTATAAAGCATTGCAAGGCTATGCCAAATCCAACGGTGTCTAGAATTAAT
AGTTGAGCCTTTTGATGATTACCAATTATTTCACTAGTTCGGGAGTTCCTAGT
AATCAGGGAAATGATAATCTGATTGAAGATTTTACCAGTTCAGGTTATCAC
CATGATGGTTTAAACAACCTGGTTTTACTGGTAAATATTTATCTTGGCACTATG
TTAAAAATTTAGAAGGTGTCACCTCTGAAACGTTACTATCTTCATTCTCTAA
GACAGGACGAGCTTTGGTTAAGAAAGCAATGTCTTTTGGAATCAAGGTTT
GCGTTCTTAAACGTGATGAGCTACATTTATTTAAAGAGATAACAACCTTCTA
CGTCAAATAGACGTGATTATATGGATAAGTCCTTAGATTATTATCAAGATT
TTTACGATAGCTTTGAAGGCAAGGCTGAATTTGTGATTGCCACTTTAAATT
TTAGAGAATACGACCATAACTTGCAAATAAAAGCTGAAGCATTGGAAAAT
AAGCTT

MTLRELTIEEFKEHSGNYDSQSFLQTPEMAKLLEKRGYDVRYLGYQVENKLEI
ISLSYIMPVTGGFQMKIDSGPVHSNSKYLKQFYKALQGYAKSNGVLELIVEPF
DDYQLFTSSGVPSNQGNLIEDFTSSGYHHDGLTTGFTGKYLSWHYVKNLE
GVTSETLLSSFSKTGRALVKKAMSFGIKVRVLKRDELHLFKEITTSTSNRRDY
MDKSLDYYQDFYDSFEGKAEFVIATLNFREYDHNLQIKAEALENKL

Sequence description

- A) Length: 771 bp - 257 aa (partial gene sequence)
B) This gene sequence was not identified using the LEEP system. It was identified immediately downstream of the ID-65 gene which was identified by

LEEP, during cloning and sequence analysis of the full-length ID-65 gene sequence. Sequence Characteristics:
No obvious leader peptide sequence
Orf is preceded by a potential Shine-Dalgarno sequence.

ID-174

Clone 2-17b (ID-80b)

TTGTCATTAAGTTTGGTTGCAGTGTTAAATCTTATCCCTCCTAAAATCATGG
GATCAGTTATTGATGCTATTACAACCTGGAAAATTAACAAGACCACAATTAC
TATGGAATTTATTAGGTTTGGTTTTGTCAGCTTTAGCTATGTATGGGCTGCG
TTATATTTGGCGTATGTATATTTTAGGGACTTCTTACAAATTAGGCCAAGTT
GTCAGATACCGTTTATTTGAACATTTTACAAAAATGTCTCCTTCTTTTTATC
AGAAATATCGTACAGGTGATTTAATGGCGCACGCGACCAACGACATCAAT
TCTCTAACACGTCTTGCAGGAGGAGGAGTTATGTCAGCAGTGGATGCCTCT
ATCACAGCATTAGTAACGCTTATCACCATGTTCTTTACTATTTTCGTGGCAA
ATGACATTAATTGCGGTTATCCCTTTGCCCTTAATGGCCTTAGCACTAGTA
AATTGGGGCGAAAAACCCATGAAACCTTCAAAGAATCTCAGGCAGCCCTT
TTCAGAATTAAATAATAAAGTG

MSLSLVAVLNLIPPKIMGSVIDAITTGKLTRPQLLWNLLGLVLSALAMYGLRYI
WRMYILGTSYKLGQVVRYRLFHFYTKMSPSFYQKYRTGDLMAHATNDINSLT
RLAGGGVMSAVDASITALVTLITMFFTISWQMTLIAVIPLPLMALALVNWGEK
PMKPSKNLRQPFSELNNKV

Sequence description

A) Length: 534 bp - 178 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-80 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-80 gene sequence. Sequence Characteristics:

No obvious leader peptide sequence
Orf is preceded by a potential Shine-Dalgarno sequence.

MHIETVIDFKELGKRYRFKNPTKELIADTLEQVLEVIKEVDYYQSQNYVVG
LSYEASAAFD SHFKVSQQLAGEHLAYFTVHKDCENEAFPLSYENVRLADNW
TANVSEQEYQEAIANIKGQIRQGNTYQVNYTLELSQQLCSD

Sequence description:

A] Length: 440 bp - 146 aa (partial gene sequence)
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-103 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-103 gene sequence. Shine Dalgarno sequence present upstream of ATG start codon, No apparent leader peptide sequence

ID-176

Clone 2-18/22b(b) (ID-104b)

GTGAATAATATGTTTTATCTCAAAATAGCCTGGCATAATTTAAAACATTCT
ATAGACCAGTACATACCATTCTCTTAGCCAGTTTATTACTTTATTCATTGA
CTTGTTCTACGCTACTAATCTTAATGAGTGCTGTTGGAAGAGATATGGGGA
CAGCGGCAACGGTTCTTTTCTTGGAGTGATTGTTTTGTCAATCTTTGCGGT
AGTCATGGAACATTATAGCTACAATATCTTGATGAAACAGCGTAGTAGTG

AATTTGGACTGTATAACATTTTGGGGATGAATAAACGTCAAGTTGCGCGTG
 TAGCTAGTCTAGAGCTGTTTATTATTTATATATTTCTTATTTCTATAGGAAG
 TCTGTTTAGTGCTTTTTTTGCTAAATTTATTTATTTAATTTTTGTCAACATTA
 5 TTAACATCATGCACTAAATCTTAGTTTAAGTTTATGGCCATTTATTATTTG
 TATCGTTATATTTACAGGTATTTTTCTGACTTTAGAAAGTTCCAGTTATTCGA
 CATGTTCAATTTATCATCCCCATTAAGTCTTTTTAGAAAGAAACAACAGGGA
 GAAAAAGAACCAAAAGGTAATCTTATACTTGCAATTTTAGCGTTAGTAGCT
 ATCGCCATCGCTTATACAATGGCTCTTACTTCAGGTAAAGCACCTGCATTA
 10 GCTGTTATCTATCGTTTCTTCTTTGCAGTACTTTTAGTAATTGCTGGTACTT
 ATCTTTTTTATATTAGTTTTATGACATGGTACTTAAAAAGGTTGCGTCAAAA
 CAAGCATTATTATTATAAATCTGAGCATTTTGTATCAACTTCGCAAATGAT
 TTTTCGAATGAAGCAAAATGCAGTAGGGTTAGCAAGTATCACTTTATTAGC
 TGTTATGGCTCTAGTTACTATTGCTACAACAGTCTCACTCTATTCAAATACA
 CAAAATGTTGTTACCGGACTATTTCCAAAATCAGTAAGTTTATCAATAGAT
 15 AATTCAAAAGGTGACGCGAAAAATATATTTGAAGAAAAGATTTTGAAGAA
 ACTAGGTAAGTCATCTAAGGAAGCTATCACTTATAATCAGACAATGATTTC
 GATGCCAGTTAGTCAATCAAGTGACTTAATATCACATCTA

 20 MNNMFYLKIAWHNLKHSIDQYIPFLLASLLLYSLTCSTLLILMSAVGRDMGTA
 ATVLFLGVIVLSIFAVVMEHYSYNILMKQRSSEFGLYNILGMNKRQVARVASL
 ELFIYIFLISIGSLFSAFFAKFIYLIFVNIINYHALNLSLSLWPFII CIVIFTGIFLTLE
 25 VPVIRHVHLSSPLSLFRKKQQGEKEPKGNLILAILALVAIAIAYTMALTSGKAP
 ALAVIYRFFFAVLLVIAGTYLFYISFMTWYLRRLRQNKHYYYKSEHFVSTSQM
 IFRMKQNAVGLASITLLAVMALVTIATTVSLYSNTQNVVTGLFPKSVSLSIDNS
 KGDAKNIFEKILKKLGKSSKEAITYNQTMISMPVSQSDDLISHL

 30

Sequence description:

- 35 A] Length: 1119 bp - 373 aa (partial gene sequence)
 B] This gene sequence was not identified using the LEEP system. It was
 identified upstream of the ID-104 gene which was identified by LEEP, during
 cloning and sequence analysis of the full-length ID-104 gene sequence.
 Possible Shine Dalgarno sequence present
 40 upstream of a GTG start codon. Possesses a potential
 leader peptide sequence

ID-177

Clone 2-5b (ID-112b)

5 ATGGTTGAGCCAATTATTTCAATACAAGGACTTCATAAAAGTTTTGGGAAA
AATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGT
GGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAAT
GAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTG
10 ATAGTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAAT
ATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAG
ACAAAAGCATAACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA
TGCTTATCCAGCAAGCTTATCTGGAGGACAACAACAACGGATTGCTATTGC
AAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCTACTTCA
15 GCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTA
GCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCA
CGTGAAGTAGCGGATCGTGTCATTATTATGGATGCAGGGATTATTGTTGAG
CAAGGGACCCCTAAGAAAGTATTTGAGCAGACAAAAGAAATCCGCACAAG
AGACTTCTTAAGTAAAGTATTATAA

20

MVEPIISIQGLHKSFGKNEVLKGIDLDIHQGEVVVIIGPSGSGKSTFLRTMNLE
VPTKGTVTFEGIDITDKKNDIFKMREKMGMVFQQFNLFNMTVLENITLSPIKT
KGLSKLDAQTKAYELLEKVGLKEKANAYPASLSGGQQQRIAIARGLAMNPDV
25 LLFDEPTSALDPEMVGEVLTVMQDLAKSGMTMVIVTHEMGFAREVADRVIF
MDAGIIVEQGTPKKVFEQTKEIRTRDFLSKVL*

30 Sequence description:

- A] Length: 735 bp - 244 aa (full length gene)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-112 gene which was identified by LEEP,
35 during cloning and sequence analysis of the full-length ID-112 gene sequence.
Shine-Dalgarno sequence precedes the 'ATG'
start codon. No obvious leader peptide

40 ID-178

Clone 2-5c (ID-112c)

ATGTCTCAATATCAAGAGTGGTTAGAAAACGACTCACTCGGTAAAGATATT
 AAGTCAGATTTAGAAGCTATTAAAGGAGATGAATCTGAAATTCAGGATCG
 TTTTACAAAACATTAGAATTTGGAACGGCGGGATTGAGAGGTAAACTTG
 5 GAGCAGGAACCAATCGTATGAATACTTATATGGTGGGGAAAGCAGCACAA
 GCATTAGCTAATCGATTATTGATCATGGCCCTGAAGCTATTGCACGTGGAA
 TTGCAGTTAGTTATGATGTCCCGTTATCAATCTAAGGAATTTGCAGAATTA
 ACTTGGTCCATTATGGCAGCAAATGGTATTAAAGCCTTATATTTA

MSHMNYKEIYQEWLENDLGLKDIKSDLEAIKGDSEIQDRFYKTLEFGTAGLR
 10 GKLGAAGTNRMNTYVMVGKAAQALANRLLIMALKLLHVELQLVMMSRYQSKE
 FAELTWSIMAANGIKALYL

Sequence description:

A] Length: 366 bp - 122 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was
 identified downstream of the ID-112 gene which was identified by LEEP,
 during cloning and sequence analysis of the full-length ID-112 gene sequence.
 20 Shine-Dalgarno sequence preceded the 'ATG'
 start codon. No obvious potential leader peptide sequence.

ID-179

Clone 2-5d (ID-112d)

ATGCAACCTGTAAAAGTCGATGAACCTTCTGTTGAAGAAACCATTACTATT
 TTGAAAGGTATCCAAAAAATAACGAAGATTATCATCACGTAAAATATAA
 30 TAATGATGCCATAGAAGCAGCTGCAGTACTATCTAATCGTTATATCCAAGA
 CCGCTTTTTACCTGATAAAGCAATAGACTTATTAGATGAAGCTGGTTCTAA
 AATGAACCTAACACTAAATTTTGTGATCCAAAAGAAATTGATCAACGTCT
 CATTGAAGCAGAAAATTTAAAAGCGCAAGCGACTCGTGAAGAAGATTACG
 AACGTGCAGCTTACTTCCGTGACCAGATTGCAAAATATAAAGAAATGCAG
 35 CAACAAAAGGTCGACGATCAAGATACACCTATTATTACCGAAAAACAAT
 TGAGCACATCATTGAAGAAAAAACGAATATCCCTGTTGGTGATTAAAAG
 AAAAAGAACAATCTCAATTAATTAATCTCGCAGATGACTTGAAACAGCAT
 GTGATCGGCCAGGATGACGCTGTCATTAAGATTGCAAAAGCTATTCGTCTGT
 AATCGAGTTGGTCTTGGTAGCCCAAACCGTCCTATTGGTTCCTTTTTATTG
 40 TAGGACCAACCGGTGTTGGTAAACTGAACTTTCTAAACAAGTACGAATTG
 AGCTCTTTGGTTCAGCTGATAGTATGATTCGTTTGTATGTGAGAGTACAT
 GGAAAAGCATGCTGTTGCTAAATTAAGTCGGAGCGCCTCCAGGATACGTGG
 GATACGAGGAAGCTGGACAACTAACTGAAAAGGTTTCGTTCGAAATCCTTAC
 TCGCTCATCCTTCTAGATGAAATTGAAAAAGCTCATCCCGATGTCATGCAT

ATGTTCTTGCAGGTCCTTGATGACGGTCGATTAACAGATGGACAAGGAAG
AACTGTTAGTTTTAAAGATACCATTATCATCATGACCTCAAATGCTGGTTC
TGGTAAACTGAAGCAAGTGTGGCTTTGGTGCCTCACGAGAAGGTAGGA
CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCAT
5 GCAAGC

MQPVKVDEPSVEETITILKGIQKKYEDYHHVKYNNDAIEAAAVLSNRYIQDRF
LPDKAIDLLDEAGSKMNLTLNFVDPKEIDQRLIEAENLKAQATREEDYERAAY
10 FRDQIAKYKEMQQQKVDDQDTPITEKTIEHIIIEKTNIPVGDLEKEQSQLINL
ADDLKQHVIGQDDAVIKIAKAIRNRVGLGSPNRPIGSFLFVGPTGVGKTELSK
QLAIELFGSADSMIRFDMSEYMEKHAVAKLVGAPPGYVGYEEAGQLTEKVRR
NPYSLILLDEIEKAHPDVMHMFLLQVLDDGRLTDGQGRTVSFKDTIIMTSNAGS
GKTEASVGFASREGRTNSSSVPGDPLESTCRHAS

15

Sequence description:

20

A] Length: 1070 bp ÷ 356 aa (Partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-112 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-
25 Dalgarno sequence preceded the 'ATG'
start codon. No obvious potential leader peptide
sequence.

ID-180

30

Clone 2-7b (ID-113b)

ATGAGAGGGAAGGTTATTTACGGCACAACCCTTATAGGTCTTTTTCTATTC
35 TTATTTTTCTATTTTTGGATTCCCTAAGCATCACATCGAGAGAATACATCATC
ATCGTATAAAGCAGGTAGATGCGAAGAGTGATTTAACAGGATTTAACAAACC
CATTTGCCCATATCAGCATTGATACAAAGCAACAAGTTATTCCTCTTGTT
ACAAAAGAAGGCGGAAAATATGTCAAAGCTAGGGATAATATTAATGTTGA
TATCGAATTACGGGATTCTCCAAGTAGATCACATCATTTATCAGAAAAGCC
40 GAGAATTAGGACAAAAGGGTTAATATCATATAGAGGAAATTCCTCTCGTT
ACTTTGATAAGAAGTCATTGAAAGTTAAGTTTGTTACTAATAAGTTAAAGG
AAAAGAAGCATCGATTAGCAGGAATGCCTAAAGAATCGGAGTGGGTATTG
CATGGTCCCTTTCTAGACAGAACATTATTAAGAAATTATCTGAGTTATAAT

5

10

15

20

25

30

35

40

5

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-117 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-117 gene sequence. N- and C-termini have yet to be determined

ID-182

Clone 3-8b (ID-120b)

ATGTACCATATTGAATTA AAAAAGGAAGCTTTACTACCAAGAGAACGCCT
AGTTGATTTAGGCGCAGATAGATTGAGTAATCAGGAGTTATTAGCCATTCT
CTTACGTACAGGTATTA AAGAAAAACCTGTTCTTGAAATTTCAACGCAAT
TTTAGAAAACATAAGCAGTTTAGCAGATTTTGGTCAATTATCCTTACAGGA
GTTGCAATCCATTAAAGGAATCGGTCAGGTTAAATCCGTCGAAATAAAAG
CTATGCTAGAACTAGCAAAACGGATTACAAAAGCTGAATATGATCGTAAA
GAGCAAATTTTAAGTAGTGAACAATTAGCGAGGAAAAATGATGCTCGAATT
AGGGGATAAAAAACAAGAACATTTAGTAGCTATTTATATGGATACACAAA
ATCGTATTATCGAACAGAGAACTATTTTTATTGGTACTGTACGTCGTTT CAG
TAGCAGAGCCAAGAGAAATTCTACATTATGCTTGTA AAAACATGGCAACT
TCTTTGATTATTATACATAATCATCCCTCAGGTTCTCCAAATCCCAGTGAAA
GTGATTTAAGTTTCACTAAAAAAATAAAACGATCATGTGATCATCTGGGAA
TTGTCTGCCTAGATCACATCATCGTTGGAAAAAATAAATATTATAGTTTT C
GAGAAGAAGCAGATATTTTATAA

MYHIELKKEALLPRERLVDLGADRLSNQELLAILLRTGIKEKPVLEISTQILENI
SSLADFGQLSLQELQSIKGIGQVKSVEIKAMLELAKRIHKA EYDRKEQILSSEQ
LARKMMLELGDKKQEHLVAIYMDTQNRHIEQRTIFIGTVRRSVAEPREILHYAC
KNMATSLIIHNPSPNPSESDLSFTKKIKRSCDHLGIVCLDHIIVGKNKYYSF
REEDIL*

Sequence description:

40

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-120 gene which was identified by LEEP.

during cloning and sequence analysis of the full-length ID-120 gene sequence.
ATG start codon is preceded by an typical
Shine-Dalgarno sequence. No obvious leader
peptide sequence

5

ID-183

10

Clone 3-11b (ID-121b)

15

TGGTTAAAAGTAGTGATAGCTTGTATTCCATCTATTTTAATTGCTTTACCAT
TTGATAATTGGTTTGAAGCTCATTTTAATTTCATGATTCCGATTGCAATAGC
CCTAATCTTTTATGGTTTTGTCTTCATATGGGTTGAAAAACGTAATGCACAC
CTCAAACACAGGTAACCGAATTGGCAAGTATGTCCTTACAAGACAGCTTTC
TTGATTGGATGTTTCCAGGTTCTCAGTATTGTTCCGGGAACCAGTCGTTCTG
GAGCTACTATTTTAGGAGCAATTATTATTGGAAGTCTGTTCCGGTCGCTG
20 CTGACTTTACTTTCTTCCTTGCCATCCCAACTATGTTTGGTTATAGTGGACT
TAAGGCGGTAAATATTTTTTAGATGGTAACGTCTTGAGTTTAGACCAATC
TTTAATACTTTTAGTAGCAAGTCTGACAGCTTTCGTAGTTAGTTTATATGTT
ATTCGTTTCTTGACAGACTATGTCAAACGACACGATTTCAACCATCTTTGGT
AAGTATCGTATAGTCTTAGGAAGTTTACTCATCCTCTACTGGTTAGTTGTTT
25 ATTTATTCTAA

30

WLKVVIACIPSILIALPFDNWFEAHFNFMIPAIALIFYGFVFIWVEKRNAHLKP
QVTELASMSYKTAFLIGCFQVLSIVPGTSRSGATILGAIIGTSRSVAADFTFFLA
IPTMFGYSGLKAVKYFLDGNVLSLDQSLILLVASLTAFVVSPLYVIRFLTDYVVKR
HDFTIFGKYRIVLGSLLILYWLVVHLF*

Sequence description:

35

A] Length: 579 bp - 193 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-68 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-68 gene sequence
described in WO 00/06736. N-terminus has yet to be determined.

40

ID-184

Clone 3-11c (ID-121c)

ATGGAAATGAAACAAATCAGTGAAACAACACTGAAAATTACAATTAGTAT
5 GGAAGATTTAGAAGATCGTGGTATGGAGCTGAAAGATTCCTAATCCCTCA
GGAGAAGACTGAGGAATTTTCTATTCTGTCATGGATGAATTAGACTTGCC
AGAAAACCTTTAAAAATAGTGGTATGTTAAGTTTTCGAGTAACACCTAAAA
AAGATCGCATTGATGTTTTTGTACAAAGTCTGAATTAAGTAAAGATTTAA
ATTTAGAAGAATTAGCAGATTTGGGTGACATTTCAAAAATGTCTCCAGAAG
10 ACTTTTTTAAAACCTTGGAACAATCGATGTTGGAAAAAGGGGATACGGAT
GCCCATGCCAAATTAGCAGAAATTGAAAATATGATGGATAAAGCAACTCA
AGAAGTAGTTGAGGAAAATGTTTCTGAAGAACAACCTGAAAAGGAAGTAG
AAACGATTGGATATGTTCACTATGTCTTTGATTTTGATAATATTGAAGCTGT
AGTTCGATTTTCACAAACGATTGATTTTCCAATAGAAGCTT
15 MEMKQISETTLKITISMEDLEDRGMELKDFLIPQEKTEEFFYSVMDELDLPENF
KNSGMLSFRVTPKKDRIDVFVTKSELSKDLNLEELADLGDISKMSPEDFFKTLE
QSMLEKGD TDAHAKLAEIENMMDKATQEVVEENVSEEQPEKEVETIGYVHY
VFDFDNIEAVVRFSQTIDFPIEA

20

Sequence description:

A] Length: 547 bp - 182 aa (Partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
25 identified downstream of the ID-68 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-68 gene sequence.
ATG start codon is preceded by an typical
Shine-Dalgarno sequence. No obvious potential
30 leader peptide
sequence

35 ID-185

Clone 3-16b (ID-122b)

40 GGAAACCAACGGCCAGTACAATCGTCAAGGGTAGATTATCCTAAACGTAG
TCGTGCCAAGATTGTAGAAGTTTATTTTAGACAAGCTTCTACTACTGATTA
TTCTGGTGTTTACAAAGGTACTATATTGACTTTGAAGCCAAAGAAACCCG
GCAGAAAACCTGCTATGCCTATGAAAAATTTTCATGCTCACCAAATAGAGC
ACATGGCAAATGTATTACAGCAAAAAGGGATTGCTTTGTCTTGCTTCATT

TTTCCACACTTAAGGAAACCTATCTACTCCCTGCTAATGAGTTAATTTTCATT
TTATCAGATTGATAAAGGCAATAAATCAATGCCTATTGATTATATCAGAAA
AAATGGATTTTTCGTAAAGGAGAGTGCCTTTCCTCAAGTCCCTTACTTAGA
TATTATTGAAGAAAAATTATTAGGCGGTGATTACAATTAA

5

GNQRPVQSSRVDYPKRSRAKIVEVYFRQASTTDYSGVYKGYIDFEAKETRQ
KTAMPMKNFHAHQIEHMANVLQQKGICFVLLHFSTLKETYLLPANELISFYQI
DKGNKSMPIDYIRKNGFFVKESAFPQVPYLDIIEEKLLGGDYN*

10

Sequence description:

A] Length: 447 bp - 149 aa (partial sequence)

15

B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-122 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-122 gene sequence. N-
terminus has yet to be determined

20

ID-186

Clone 3-17b (ID-123b)

25

GGATCCTAAAAACGCTAAGGTTTATCAAAAAAATGCTGATCAATTTAGTG
ACAAGGCAATGGCTATTGCAGAGAAGTATAAGCCAAAATTTAAAGCTGCA
AAGTCTAAATACTTTGTGACTTCACATACAGCATTCTCATACTTAGCTAAG
30 CGATACGGATTGACTCAGTTAGGTATTGCAGGTGTCTCAACCGAGCAAGA
ACCTAGTGCTAAAAAATTAGCCGAAATTCAGGAGTTTGTGAAAACATATA
AGGTTAAGACTATTTTTGTTGAAGAAGGAGTCTCACCTAAATTAGCTCAAG
CAGTAGCTTCAGCTACTCGAGTTAAAATTGCAAGTTTAAGTCCTTTAGAAG
CAGTTCCCAAAAACAATAAAGATTACTTAGAAAATTTGGAAACTAATCTTA
35 AGGTACTTGTCAAATCGTTAAATCAATAG

DPKNAKVYQKNADQFSDKAMAIAEKYKPKFKAAKSKYFVTSHTAFSYLAKR
YGLTQLGIAGVSTEQEPSAKKLAEIQEFVKTYKVKTIFVEEGVSPKLAQAVAS
40 ATRVKIASLSPLEAVPKNNKDYLENLETNLKVLVKS LNQ*

Sequence description:

A] Length: 433 bp - 144 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-123 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-123 gene sequence. N-terminus has yet to be determined

ID-187

Clone 3-46/47 (ID-130b)

15 ATGAAAAAAGTCATCGATT TAAAAAACTACAAAAAGCATACGCCTCAGA
AACTGTTTTAAATAATATTAATTTGGAGGTGTTTAAAGGAGAAATAATTGG
ATTAATAGGACCCTCTGGAGCAGGGAAATCTACCTTGATTAAAACTATGCT
TGGCATGGAAAAAGCAGATAAGGGAACAGCTCTTGTTCTTGATACTCAAA
20 TGCCAGATCGTAATATTTTAAATCAAATTGGCTATATGGCTCAATCTGATG
CCTTACACGAGTCTTTAACTGGCTTAGAAAATTTATTATTCTTTGGAAAAA
TGAAAGGTATTCAAAAAACTGAATTA AAAACAGCAGATAACTCATATTTCT
AAAGTAGTAGATCTAGAAAACCAACTTGATAAATTTGTCTCAGGTTACTCA
GAAGGTATGAAAAGACGGCTTTCTCTAGCCATCGCCCTACTTGGAACCCC
25 ACAGTTTTAATCCTAGATGAACCTACCGTTGGAATTGATCCATCCTTGAGG
AGAAAAATCTGGCAAGAGCTAATTAATATTAAGGATGAAGGACGTTCTAT
CTTTATTACAACCCACGTTATGGATGAAGCAGAATTAACAAGTAAGGTTGC
ACTACTATTACGTGGAAACATTATTGCCTTTGATACTCCATTACATTTAAA
AAAACAATTTAATGTGAGTACTATTGAGGAAGTTTTCTTAAAGCTGAAGG
30 AGAATAA
MKKVIDLKKLQKAYASETVLNNINLEVFKEIIGLIGPSGAGKSTLIKTMLGME
KADKGTALVLD TQMPDRN ILNQIGYMAQSDALHESLTGLENLLFFGKMKGIQ
KTELKQQITHISKVVDLENQLDKFVSGYSEG MKRRLSLAIAL LGNPTVLILDEP
TVGIDPSLRRKIWQELINIKDEGRSIFITTHVMDEAELTSKVALLLRGNIIAFDTP
35 LHLKKQFNVSTIEEVFLKAEGE*

Sequence description:

A] Length: 717 bp - 239 aa (Possible full-length sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-130 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-130 gene sequence. ATG start codon is preceded by a possible

Shine-Dalgarno. No obvious potential leader
peptide sequence

5

ID-188

10

Clone 3-83b (ID-144b)

15

ATGGTACAAATGATACATGATATGATTAAAACAATTGAGCATTTTGCTGAG
ACACAAGCTGATTTTCCAGTGTATGATATTTTAGGGGAAGTCCATACTTAT
GGACAACCTTAAAGTAGACTCTGACTCTCTAGCTGCTCATATTGATAGCCTA
GGCCTTGTTGAAAAATCACCTGTCCTTAGTATTCGGTGGTCAAGAATATGAA
ATGTTGGCGACATTTGTTGCTTTAACAAAGTCAGGGCATGCTTATATACCG
GTTGACCAACACTCTGCTTTGGATAGAATACAGGCTATTATGACAGTTGCT
CAACCAAGCCTTATCATTTC AATTGGTGAATTCCTCTTGAAGTTGATAAT
GTCCCAATCCTAGACGTTTCTCAAGTTTCAGCTATTTTGAAGAAAAGACT
CCTTATGAGGTAACACATTCTGTAAAGGTGATGATAATTACTATATTATT
TTCACCTCAGGGACTACTGGTTTACCAAAAAGGTGTGCAAATTTACATGAC
AATTTATTGAGCTTTACAAATTGGATGATTTCTGATGATGAGTTTTCAGTTC
CTGAAAGACCGCAAATGTTGGCTCAACCC

25

MVQMIHDMIKTIEHFAETQADFPVYDILGEVHTYGQLKVDSDSLAAHIDSLGL
VEKSPVLVFGGQEYEMLATFVALTKSGHAYIPVDQHSALDRIQAIMTVAQPSL
IISIGEFPLEVDNVPILDVSQVSAIFEEKTPYEVTHSVKGDDNYIIFTSGTTGLP
KGVQISHDNLLSFTNWMISDDEFSVPERPQMLAQP

30

Sequence description:

35

A] Length: 592 bp - 197 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-144 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-144 gene sequence.
Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No obvious
leader peptide sequence
This orf is not in frame with nuc

40

ID-189

5

10

15

20

25

30

35

40

Sequence description:

A] Length: 1126 bp - 393 aa (partial gene

sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-145 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-145 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a possible leader peptide sequence.

ID-190

Clone 3-94b

TCAGAAAATGCAGAGGCAGCAACGGTTGCCACAAACTTGGTTACCAAAGG
AGCTAATGTCATTATCGGACCAGCAACATCGGGTGCAGCTGCATCTTCAAC
TCCAAAAGTAAATGCAGCAGCAGTTCCAATGATTGCACCTGCTGCGACAC
AAGACAATTTAGTCTATGGTTCTGATGGAAAAACCTTAAATCAGTATTTCT
TCCGAGCTACTTTTGTCTGATAATTATCAAGGAAAGCTATTGTCTCAGTATG
CTACAGACAACCTTAAAGCTAAAAAAGTTGTTCTATTTTATGATAATTCAT
CAGATTACTCAAAGGGGGTAGCAAAATCATTAAAGGAAAGTTATAGTGGA
AAAATTGTTGATAGTATGACATTCTCCGCTGGTGATACTGATTTCOAAGCG
TCATTGACTAAGTTGAAAGGGAAAGAATATGATGCTATTGTGATGCCAGG
TTACTATACCGAGACAGGATTAATAGTTAAGCAAGCGCGTGATTTAGGTAT
CTCTAAACCGGTTCTTGGGCCTGATGGTTTTGATAGTCCGAAATTTGTGCA
ATCGGCAACACCTGTAGGAGCTTCAAACGTTTATTATTTGACAGGTTTCAC
TACACAAGGATCAACCAAAGCTAAAGCT

SENAEAATVATNLVTKGANVIIGPATSGAAASSTPKVNAAAVPMIAPAATQD
NLVYGSDGKTLNQYFFRATFVDNYQGKLLSQYATDNLKAKKVVLFDNSSD
YSKGVAKSFKESYSGKIVDSMTFSAGDTHFQASLTKLKGKEYDAIVMPGYT
ETGLIVKQARDLGISKPVLGPDGFDSPKFVQSATPVGASNVYYLTGFTTQGST
KAKA

Sequence description

A] Length: 637 bp - 231 aa (partial sequence)

5

10

20

25

30

35

40

ATGACTAATATCTCAGATGTTCCAAAAGCTATTAGAACACAGGCACAGTAT
GTTCTCTTGGGAATGAGAGTTATGGATCAGTCGGTATTACCGAAAACATAT
AATTCAAAAAGAACCTTATTTGAAACCAGATATGATTTATATTCATGATAGA

AGACAAGAGACAATGCTTAAAATCACTCAAGAAATAGAAATGGAGCATTG
A

MTNISDVPKAIRTQAQYVLLGMRVMDQSVLPKTYNSKEPYLKPDMIYIHDRR
QETMLKITQEIEMEH*

Sequence description

A] Length: 204 bp - 68 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-155 gene which was identified by LEEP, during cloning and sequence analysis of the ID-155 gene sequence.

ATG start codon is preceded by a potential typical Shine-Dalgarno sequence.

Has a

typical leader peptide. N-terminus has yet to be verified

ID-193

Clone 2-54altb (ID-172b)

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCTGGGGGAATATAAATT
TGGATTTTCATGACGATGTAAAGCCAATTTATTCTACGGGAAAAGGTCTAAA
TGAGGCTGTTATTCGTGAGTTATCTGCAGCTAAGGGTGAACCTGAGTGGAT
GTTGGACTTTCGTCTAAAATCCTTGGAACGTTTAATAAAAATGCCGATGCA
GACCTGGGGAGCAGATTTATCAGATATTGATTTTGATGATATTATTTATTA
TCAAAAAGCATCTGATAAACCTGCGCGTGATTGGGATGATGTTCCAGAAA
AAATCAAAGAACTTTTGAAAGAATTGGGATTCCAGAAGCTGAAAGAGCC
TATCTTGCAGGAGCATCAGCACAATATGAATCAGAAGTAGTTTATCACAAT
ATGAAAGAAGAATATGATAAGCTGGGTATTGTTTTTACGGATACTGACTCT
GCACTTAAAGAGTACCCAGAGCTATTCAAAAAATATTTTGCTAAACTTGTC
CCTCCAACAGATAATAAATTAGCTGCTCTGAACTCTGCTGTATGGTCAGGT
GGAACATTTATTTATGTTCCCTAAAGGTGTAAAGGTGGATATTCCACTTCAA
ACTTACTTCCGTATTAATAATGAAAATACTGGACAATTTGAACGTACTCTC
ATTATTGTTGATGAGGGAGCAAGTGTTCACTATGTTGAAGGTTGTACCGCC
CCAACCTTATTCTTCAAATAGTTTACATGCAGCTATAGTTGAAATTTTTGCAC
TTGATGGAGCTTATATGCGCTATACGACTATTCAAAATTTGGTCCGATAATG
TCTATAATTTAGTGACAAAACGTGCTACCGCTAAAAAAGATGCAACAGTT
GAGTGGATAGATGGAAATCTAGGAGCTAAAACAACAATGAAATACCCATC

5

15

25

- 30

35

40

ATGATAGAATTCTTTTCTAATATCAGAACAGAGATTCCGCAGATGCCTTTA
CTTATCCATAGTTTGATTTTATCTGTCTTACCTTTTCTGATGTGGCTGACTTT
GGTTAATAGAGATAAGCCTTTGTATAAACTATTTGGAGTATCCTTTTAGG
ACTTCAGTTAATTACGATTTATACTTGGTTTTTCTGGGCAAAATTGCCTTTA

TCTGAAAGTCTTCCCCTTTACCATTGTCGAATAGGCATGTTTGTCCGGTCTCT
TA

5 MIEFFSNIRTEIPQMPLLIHSLILSVLPFLMWLTLVNRDKPLYKTIWSILLGLQLI
TIYTWFFWAKLPLSESLPLYHCRIGMFVGLL

Sequence description

- 10 A) Length: 261 bp - 87 aa (partial gene sequence)
B) This gene sequence was not identified using the LEEP system. It was identified
downstream of the ID-81 gene which was identified by LEEP, during cloning and
sequence analysis of the full-length ID-81 gene sequence. Sequence Characteristics:
Possesses a potential leader peptide sequence. Orf is preceded by a potential Shine-
Dalgarno sequence.

15

ID-195

20

Clone RS-55b

25 AAGCTTGTGCAAAGTATTAAAGAGATAGGATTAGCTAATGCGCATTTATTA
GCTGTTGCTCCGACAGGGTCAATCAGTTATCTTTCTTCTTGTACTCCGAGCC
TTCAACCGGTTGTATCACCTGTCGAAGTACGCAAGGAAGGAGCACTGGGG
AGGGTTTATGTAGCTGCTTATAAGATTGATGCAGATAATTATGTCTACTAC
AAAAAAGGAGCTTATGAAGTGGGATCTGAGGCGATTATCAATATTGCAGC
TGCCGCTCAAAAACACATTGATCAAGCTATTTTCGTTAACGCTTTTCATGAC
30 AGATCAAGCAACTACGCGAGATTTAAATAAAGCCTATATTCAAGCATTTA
AACAAAAATGTGCCTCTATTTATTATGTACGAGTGAGACAGGACATCCTAG
AAGGTAGCGAGAGTTATGATGATATGCTGGATGATTTCACTTCATCGGACT
TAGAAGACTGTCAATCCTGCATGATTTAA

35

>KLVQSIKEIGLANAHLAVAPTGSISYLSSCTPSLQPVVSPVEVRKEGALGRV
YVAAYKIDADNYVYYKKGAYEVGSEAIIINIAAAQKHIDQAISLTLFMTDQAT
TRDLNKAYIQAFKQKCASIYYVRVRQDILEGSESYDDMLDDFTSSDLEDCQSC
MI*

40

Sequence description:

A] Length 486 bp - 162 aa (Partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-87 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-87 gene sequence. N-terminus to be determined.

ID-196

Clone RS-59(ID-90b)

GTGAGGACATATATTACAAACTTGAATGGACATTCAATCACTAGTACAGC
ACAAATAGCTCAAAACATGGTAACAGATATAGCAGTAAGCTTAGGTTTTTC
GTGAGCTGGGAATACATTCTTATCCGATTGATACTGATTCTCCTGAGGAAA
TGAGTAAGCGTTTAGATGGAATCTGTTCCGGACTTAGAAAAAATGATATTG
TCATATTTTCAGACACCTACATGGAACACTACAACCTTTTGATGAAAAATTAT
TTCACAAATTAAAAATATTTGGTGTAAAGATTGTTATTTTTATACATGATGT
TGTACCGCTAATGTTTGATGGAAATTTTTATTTGATGGATAGAAGTATAGC
TTATTATAATGAAGCAGATGTTTAATAGCCCCTAGTCAAGCAATGGTTCGAT
AAGCTT

MRTYITNLNGHSITSTAQIAQNMVTDIAVSLGFRELGIHSYPIDTDSPEEMSKRL
DGICSGLRKNDIVIFQTPTWNTTTFDEKLFHKLKIFGVKIVIFIHDVVPLMFDGN
FYLMDRTIAYYNEADVLIAPSQAMVDKL

Sequence description:

A] Length: 414 bp - 138 aa(partial gene)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence.

No obvious signal peptide, but a possible Shine Dalgarno sequence is present upstream of ATG start codon. C-terminus has yet to be determined.

ID-197

Clone RS-59c (ID-90c)

5 CATGGAAATGAAGTTGATGATGTTATTAGAAGGGCATTGGAATATAATCAC
CTTATCTTTGCTTTTGATAATACCTGTCATAACAGAGAGTTAGTATTAGATA
GCAATATCATTCTCACACAACCTGTGAACAATTGATAAATTTAATGAAAA
ATTTATCAGGCTCCATTATGTATTTGCTAGAGCAACAAAGAGAACAAACA
AGTAATGAAACAAAAGAGCGTTATAAAGAAATATTAGGAGGGTATGGAA
ATGCCTAA

10 HGNEVDDVIRRAFEYNHLIFAFDNTCHNRELVLDSNIISHTTCEQLINLMKNLS
GSIMYLLEQQREQTSNETKERYKEILGGYGNA*

15 Sequence description:

A] Length: 261 bp - 87 aa(partial gene sequence)

20 B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-90 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-90 gene sequence. N-
terminus has yet to be determined

25 ID-198

Clone RS-70b (ID-93b)

30 ACATTTTTATATTATGTATTTGAAGACGTAGCCACCCAGTCAAATATGACT
GGGAAGATTTTTAGTATGTCTAAAGAAGAGTTGTCATATTTACCCGTTATT
AAACTTTTTAAGAATCAAGGTGTATACAACGGCTTGATTGGTCTATTCCTC
CTTTATGGGTTATATATTTACACAGAATCAAGAAATTGTAGCTATTTTTTTAA
TCAATGTGTTGCTAGTTGCTGTTTATGGTGCTTTGACAGTTGATAAAAAAA
TCTTATTAAACAGGGTGGTTTACCTATATTAGCTCTTTTAACATTCTTATT
35 TTAA

TFLYYVFEDVATQSNMTGKIFSMSKEELSYLPVIKLFKNQGVYNGLIGLFLLY
GLYISQNQEIVAIFLINVLLVAVYGALTVDKKILLKQGGLPILALLTFLF*

40

Sequence description:

A] Length: 312 bp - 104 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence.

5 N-terminus has yet to be determined

ID-199

10 Clone RS-70c (ID-93c)

15 ATGAAATTAAGTGTCTTGATTATGGGCTTATTGATTATGGAAAACTGCA
AGTGATGCAATACAAGAAACGATTCTTTTATCACAAGAGGCGGAGCAACT
AGGCTATCATCAATTTTGGGTGGCTGAACATCACGGTGTTAAGGCATTTCAG
TATTAGCAATCCAGAATTAATGATAATGCATTTGGCTAACCAGACTAAATC
TATCAAAATTGGCTCTGGAGGTATAATGCCTCTGCACTATAGTAGTTTTAA
ACTCGCGGAGACTCTCAAGACATTAGAGACATGTCATCCGAATCGAGTAA
GTATTGGTTTATAGGAAATTCAGTGGGACAGTTAAAGTTTCAAATGCACTTC
GTAGCTTACATAAAGCACATGATTACGAAGAGGTACTGGAGGAATTGAAG
20 TCATGGCTTATTGATGAATCATCCAGTAAGGAACCATTAGTTCAACCGACT
CTTTCTAGCTTCCCAGACTTATATGTGTTGGGGAGTGGTCAAAAATCAGCT
TATTTAGCGGCTAAACTTGGCTTAGGCTTACCTTCGGTGTTCCTTTTA
TGGACAAAGACCCATTGACAGAAGCTAAA

25 MKLSVLDYGLIDYGKTASDAIQETILLSQEAEQLGYHQFWVAEHHGVKAFFSIS
NPELMIMHLANQTKSIKIGSGGIMPLHYSSFKLAETLKTLETCHPNRVSIGLGN
SLGTVKVSNALRSLHKAHDYEEVLEELKSWLIDESSSKEPLVQPTLSSFPDLYV
LGSGQKSAYLAAKLGLGFTFGVFPFMDKDPLEAK

30 Sequence description:

A] Length: 588 bp - 196 aa (partial)

35 B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence. No obvious signal peptide, but Shine Dalgarno sequence upstream of the ATG start codon.